Campbell Systematic Reviews
2011:8
First published: 14 November, 2011
Last updated: 14 November, 2011
Search date: April, 2011

Dropout prevention and intervention programs: Effects on school completion and dropout among school-aged children and youth

Sandra Jo Wilson, Emily E. Tanner-Smith, Mark W. Lipsey, Katarzyna Steinka-Fry, & Jan Morrison

THE CAMPBELL COLLABORATION

Colophon

Title	Dropout prevention and intervention programs: Effects on school completion and dropout among school-aged children and youth
Institution	The Campbell Collaboration
Authors	Wilson, Sandra Jo Tanner-Smith, Emily E. Lipsey, Mark W. Steinka-Fry, Katarzyna Morrison, Jan
DOI	10.4073/csr.2011.8
No. of pages	62
Last updated	24 August, 2011
Citation	Wilson SJ, Tanner-Smith EE, Lipsey, MW, Steinka-Fry, K, Morrison, J. Dropout prevention and intervention programs: Effects on school completion and dropout among school aged children and youth. Campbell Systematic Reviews 2011:8 DOI: 10.4073/csr.2011.8

Copyright	© Wilson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords	School dropout, school attendance, early school leaving, school failure
Contributions	Wilson, Tanner-Smith, and Lipsey contributed to the writing and revising of this review and protocol. Wilson, Tanner-Smith, Steinka-Fry and Morrison contributed to information retrieval and data collection.
Support/Funding	Work on this review was supported by the Campbell Collaboration and the Peabody Research Institute at Vanderbilt University.
Potential Conflicts of Interest	The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

Corresponding author	Sandra Jo Wilson Peabody Research Institute, Vanderbilt University 230 Appleton Place, PMB 181 Nashville, TN 37203-5721 USA E-mail: sandra.j.wilson@vanderbilt.edu

Campbell Systematic Reviews

Editors-in-Chief	Mark W. Lipsey, Vanderbilt University, USA Arild Bjørndal, Norwegian Knowledge Centre for the Health Services & University of Oslo, Norway
Editors	
Crime and Justice	David B. Wilson, George Mason University, USA
Education	Sandra Jo Wilson, Vanderbilt University, USA
Social Welfare	William Turner, University of Bristol, UK
Managing Editor	Karianne Thune Hammerstrøm, The Campbell Collaboration
Editorial Board	
Crime and Justice	David Weisburd, Hebrew University, Israel & George Mason University, USA Peter Grabosky, Australian National University, Australia
Education	Gary W. Ritter, University of Arkansas, USA Paul Connolly, Queen's University Belfast, UK
Social Welfare	Aron Shlonsky, University of Toronto, Canada Paul Montgomery, University of Oxford, England
Methods	Therese Pigott, Loyola University, USA Peter Tugwell, University of Ottawa, Canada

The Campbell Collaboration (C2) was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. C2 offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of C2's editors, librarians, methodologists and external peer-reviewers contribute.

The Campbell Collaboration
P.O. Box 7004 St. Olavs plass
0130 Oslo, Norway
www.campbellcollaboration.org

Table of contents

TABLE OF CONTENTS **4**

LIST OF TABLES **6**

LIST OF FIGURES **7**

EXECUTIVE SUMMARY/ABSTRACT **8**
Background 8
Objectives 8
Search Strategy 8
Selection Criteria 8
Data Collection and Analysis 9
Results 9
Author's Conclusions 10

1 INTRODUCTION **11**
1.1 Rationale 11
1.2 Dropout Interventions 13
1.3 Previous Reviews on Dropout Interventions 14
1.4 Objectives of the Review 16

2 METHODS **17**
2.1 Title Registration and Review Protocol 17
2.2 Eligibility Criteria 17
2.3 Literature Search 20
2.4 Analytic Methods 23

3 RESULTS **25**
3.1 Literature Search & Study Identification 25
3.2 Selecting Studies for Analysis 26
3.3 General Dropout Programs 26
3.4 Dropout Programs for Teen Parents 38
3.5 Publication Bias Analysis 45
3.6 Sensitivity Analysis 48

4 DISCUSSION **49**
4.1 Summary of Evidence 49

4.2 Limitations 51
4.3 Conclusions 52

5 REFERENCES 54

6 APPENDIX I: ADDITIONAL TABLES 56
6.1 Number of Reports Excluded, By Specific Exclusion Criteria 56
6.2 Bivariate Correlation Matrix of Study Characteristics—General Programs 58
6.3 Bivariate Correlation Matrix of Study Characteristics—Teen Parent Programs with Dropout or Graduation Outcomes 60

7 APPENDIX II: ADDITIONAL FIGURES 61
7.1 Study Identification Flow Diagram 61
7.2 Forest Plot for Teen Parent Programs Providing School Enrollment Outcomes 62

List of Tables

Table 3.1:	Intervention Types – General Dropout Programs	27
Table 3.2:	Characteristics of Studies of General Dropout Programs	29
Table 3.3:	Mean Odds Ratio Effect Sizes and Heterogeneity Statistics for General Dropout Programs	32
Table 3.4:	Meta-Regression Results Predicting Logged Odds Ratio Effect Sizes – General Dropout Programs	34
Table 3.5:	Mean Adjusted Odds Ratios and Success Rates by Intervention Type	38
Table 3.6:	Intervention Types – Teen Parent Programs	39
Table 3.7:	Characteristics of Included Studies – Teen Parent Programs	39
Table 3.8:	Mean Odds Ratio Effect Sizes and Heterogeneity Statistics by Intervention Type	41
Table 3.9:	Meta-Regression Results Predicting Logged Odds Ratio Effect Sizes – Teen Parent Programs	43
Table 3.10:	Odds Ratios in Percentage Form	43
Table 3.11:	Results of Peters' Regression Test for Small Study Effects By Type of Program	48

List of Figures

Figure 1.1: International Completion Rates for Upper Secondary Education **Error! Bookmark not define**
Figure 3.1: Mean Adjusted Odds Ratios for the General Dropout Programs 37
Figure 3.2: Mean Adjusted Odds Ratios for the Teen Parent Dropout Programs 45
Figure 3.3: Funnel Plot for the General Dropout Programs 46
Figure 3.4: Funnel Plot for the Teen Parent Dropout Programs 47

Executive Summary/Abstract

BACKGROUND

Dropping out of high school is associated with numerous detrimental consequences, including low wages, unemployment, incarceration, and poverty. There are a large number of school and community-based prevention and intervention programs for general population and at-risk students, and there are a number of programs designed specifically to encourage school completion among pregnant and parenting teens. No comprehensive systematic reviews have examined these programs' overall effectiveness.

OBJECTIVES

The objectives of this systematic review were to summarize the available evidence on the effects of prevention and intervention programs aimed at primary and secondary students for increasing school completion or reducing school dropout. The primary focus of the meta-analysis was to examine the comparative effectiveness of different programs and program approaches in an effort to identify those with the largest and most reliable effects on school completion and dropout outcomes. We also sought to summarize the effects of programs designed for pregnant and parenting teens.

SEARCH STRATEGY

A comprehensive and diverse international search strategy was used to locate qualifying studies reported between 1985 and 2010. A wide range of electronic bibliographic databases were searched, along with research registers, other grey literature databases, reference lists of all previous meta-analyses and reviews on the topic, as well as citations in research reports. We also maintained correspondence with researchers in the field of school dropout prevention.

SELECTION CRITERIA

Studies eligible for inclusion in the review were required to meet several eligibility criteria. First, studies must have involved a school-based or affiliated psychological, educational, or behavioral prevention or intervention program intended to have

beneficial effects on students recipients; or, evaluated a community-based program explicitly presented as a dropout prevention or intervention program. Second, studies must have investigated outcomes for interventions directed toward school-aged youth, or those expected to attend pre-k to 12th grade primary and secondary schools or the equivalent. Third, studies must have used experimental or quasi-experimental research designs, including random assignment, non-random assignment with matching, or non-random assignment with statistical controls or sufficient information to permit calculation of pre-treatment effect size group equivalence. Fourth, studies must have reported at least one eligible outcome variable measuring school completion, school dropout, high school graduation, or school attendance. Finally, the date of publication or reporting of the study must have been in 1985 or later.

DATA COLLECTION AND ANALYSIS

The literature search yielded a total of 23,677 reports, 2,794 which were deemed potentially relevant and retrieved for eligibility determination. Of those, 548 reports describing 167 different studies were included in the final review. Random-effects inverse variance weighted meta-analytic methods were used to synthesize odds ratios for the school dropout outcomes. Meta-regression models were used to examine the effects of program characteristics, methodological characteristics, and participant characteristics on the dropout odds ratios. Funnel plots, regression tests for small sample bias, and trim and fill analyses were used to assess the possibility of publication bias.

RESULTS

General dropout programs (152 studies; 317 independent samples) and dropout programs for teen parents (15 studies; 51 independent samples) were analyzed in separate meta-analyses. Overall, both general dropout programs and programs specialized for teen parents were effective in reducing school dropout (or increasing school completion). The random effects weighted mean odds ratio for the general programs was 1.72. Using the average dropout rate for control groups of 21.1%, the odds ratio for the general programs translates to a dropout rate of 13%. For the teen parent programs, the mean odds ratio for graduation and dropout outcomes was 1.83 and was 1.55 for school enrollment outcomes. The average graduation rate for the young women in comparison groups was 26%. The corresponding graduation rate for young mothers in the intervention programs was 39%. For school enrollment outcomes, the average enrollment rate for the comparison mothers was 33%. The mean odds ratio of 1.55 for these studies translates into an enrollment rate of about 43%. Moderator analyses for the general programs indicated that studies with similar program and comparison groups at baseline and those that provided posttest data adjusted for baseline non-equivalence produced smaller effect sizes. For teen parent programs, moderator analyses found that random and matched

designs produced smaller effect sizes than non-random or non-matched designs. Effect sizes were therefore adjusted for methodological characteristics to examine the effects of different program types net of the influence of method. The effects were generally consistent across different types of programs and for different types of participant samples. However, higher levels of implementation quality tended to be associated with larger effects. Analyses provided no strong indication of the presence of publication or small study bias.

AUTHOR'S CONCLUSIONS

Overall, results indicated that most school- and community-based programs were effective in decreasing school dropout. Given the minimal variation in effects across program types, the main conclusion from this review is that dropout prevention and intervention programs, regardless of type, will likely be effective if they are implemented well and are appropriate for the local environment. We recommend that policy makers and practitioners choosing dropout prevention programs consider the cost-effectiveness of programs, and choose those that fit best with local needs as well as implementer abilities and resources.

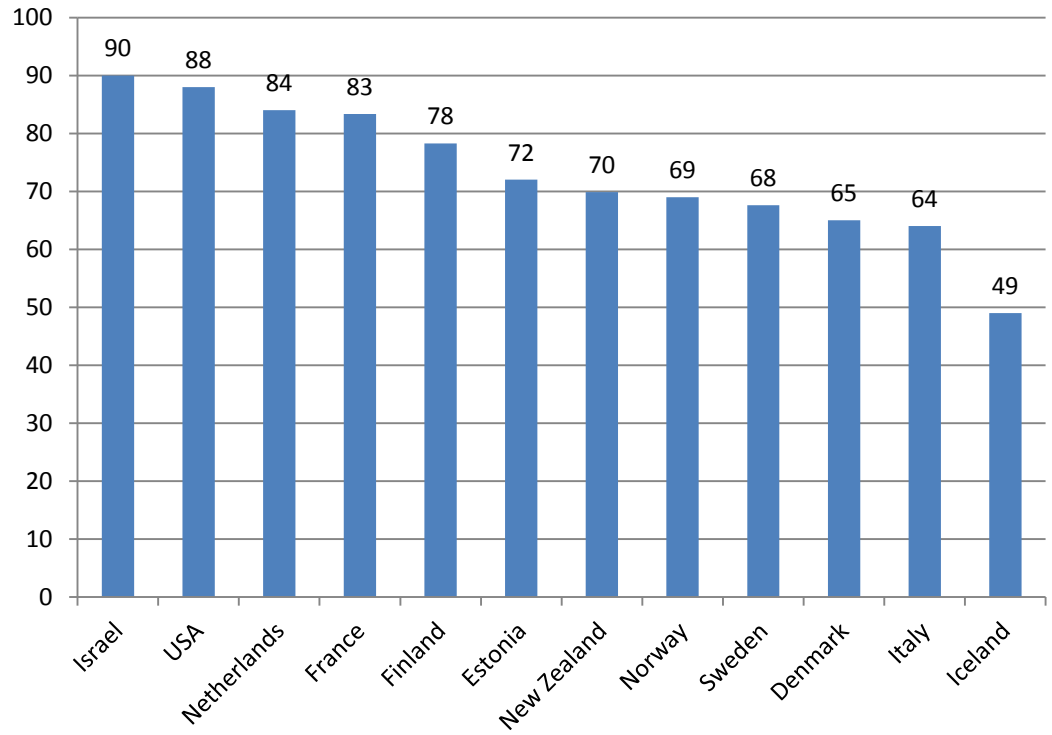
1 Introduction

1.1 RATIONALE

With the expansion of regional and national economies into a global marketplace, education has critical importance as a primary factor in allowing young adults to enter the workforce and advance economically, as well as to share in the social, health, and other benefits associated with education and productive careers. Dropping out of school before completing the normal course of secondary education greatly undermines these opportunities and is associated with adverse personal and social consequences. Dropout rates in the United States vary by calculation method, state, ethnic background, and socioeconomic status (Cataldi, Laird, & KewelRamani, 2009). Across all states, the percentage of freshman who did not graduate from high school in the expected four years ranges from 13.1% to 44.2% and averages 26.8%. The status dropout rate, which estimates the percentage of individuals in a certain age range who are not in high school and have not earned a diploma or credential, is slightly lower. In October 2007, the proportion of non-institutionalized 18-24 year olds not in school without a diploma or certificate was 8.7%. Status dropout rates are much higher for racial/ethnic minorities (21.4% for Hispanics and 8.4% for Blacks vs. 5.3% for Whites). In general, males are more likely to be dropouts than females (9.8% vs. 7.7%), but teenage pregnancy and parenthood are particularly strong risk factors for young women, especially in the United States (Dalton, Glennie, & Ingels, 2009). In fact, only about 50% of American teen mothers receive a high school diploma by age 22 (Perper, Peterson, & Manlove, 2010). Event dropout rates illustrate single year dropout rates for high school students and show that students from low-income households drop out of high school more frequently than those from more advantaged backgrounds (8.8% for low-income vs. 3.5% for middle income and 0.9% for high income students). Further, the National Dropout Prevention Center/Network reports that school dropouts in the United States earn an average of $9,245 a year less than those who complete high school, have unemployment rates almost 13 percentage points higher than high school graduates, are disproportionately represented in prison populations, are more likely to become teen parents, and more frequently live in poverty (2009). The consequences of school dropout are even worse for minority youth, further exacerbating the economic and structural disadvantages they must often face (e.g., Sum, Khatiwada, & McLaughlin, 2009).

Internationally, completion rates for secondary schooling vary widely, even across developed countries. In Canada, the proportion of 20-24 year olds who have not completed high school and are not attending school is about 9% (Canadian Council on Learning, 2010). Figure 1.1 below presents recent secondary school completion rates for 12 countries (see also Arneberg, 2009). Completion rates are lowest in Iceland (around 49%) and highest in Israel (about 90%). Across Europe, rates vary from country to country.

Figure 1.1: International Completion Rates for Upper Secondary Education



Notes: Denmark, Finland, Netherlands: Students who started upper secondary education in 2001, completion measured in 2006; Norway, Sweden: Students who started upper secondary in 2002, completion measured in 2007; France: Students who started lower secondary education in 1995, completion measured in 2007 includes those still in education; Iceland: Preliminary figures based on persons born in 1982 (most started upper secondary education in 1998), graduation measured in 2003; Italy: Persons born in 1987 (most entered ISCED 3 in 2001-02), graduation measured in 2005-06 . New Zealand: 9th grade cohort of 2001, (entering ISCED 3 in 2003), completions status measured at end of 2007; USA 10th graders in 2002, completion status measured in 2006. Israel: Students who started upper secondary in 2004, completion measured in 2007 (after 3 years). Estonia: Students who started upper secondary education in 2003, completion status measured in 2008.
Figure provided by Marie Wenstøp Arneberg, Senior Advisor of the Norwegian Ministry of Education and Research.

School dropout has implications not only for the lives and opportunities of those who experience it, but also has enormous economic and social implications for society at large. For instance, recent estimates suggest that relative to high school graduates, the average U.S. student who drops out of high school costs taxpayers over $292,000 due to lower tax revenues, higher cash and in-kind transfer costs, and costs associated with incarceration (Sum, Khatiwada, & McLaughlin, 2009). The

Organisation for Economic Co-operation and Development (2009) has similarly documented the tremendous social and economic gains associated with secondary school completion in OECD member countries.

1.2 DROPOUT INTERVENTIONS

A relatively large number of intervention and prevention programs in the research literature give some attention to reducing dropout rates. The National Dropout Prevention Center/Network, for instance, currently lists hundreds of "model programs." Relatively few of those programs, however, bill themselves as dropout programs; many focus on academic performance, risk factors for dropout such as absences or truancy, or indirect outcomes like student engagement, but may also include dropout reduction as a program objective. The corresponding research domain includes evaluations of virtually any program provided to students for which dropout rates are measured as an outcome variable, regardless of whether the programs are explicitly billed as dropout programs. To represent the full scope of relevant research on this topic, all such programs should be considered in a review of dropout programs. We turn now to a brief description of some of the more common program types in the literature.

Many programs that give attention to reducing school dropout, whether specifically billed as dropout programs or not, are informed by research on the risk factors associated with school dropout (cf. Hammond, Linton, Smink, & Drew, 2007). School or class restructuring programs are common in the literature on school dropout, and are becoming increasingly popular in U.S. secondary schools. The simplest form of these programs involves creating smaller classes or lower student-teacher ratios. The more complex school restructuring programs involve larger scale school restructuring activities that may include some combination of smaller classes, blocked schedules, personalized learning settings in which students spend more time with fewer teachers, and/or "academies" focused around grade level (e.g., 9th grade academies) or career foci. The general rationale behind school restructuring programs is that by creating more personalized and relevant instruction, students will be more focused and engaged in school. Another prevalent type of dropout program focuses on vocational training, work-related coursework, career exploration, internships, and/or paid employment for students. As with the school restructuring programs, vocationally-oriented programs seek to create more relevant educational experiences for a broad range of students. Supplemental academic services (e.g., tutoring, homework assistance, remedial education) are an obvious strategy for preventing dropout, and are common in educational settings throughout the United States and Europe. As we shall see below, there are a variety of other programs for preventing dropout, including community service programs, college-oriented programming, case management, counseling, skills training, as well as multi-service packages that offer a range of diverse services.

In addition, teenage pregnancy and parenthood are among the biggest risk factors for dropping out of high school for American girls. As a result, a variety of programs designed to encourage young mothers to complete schooling are available, and many involve contingencies in which welfare payments or other incentives are tied to school enrollment and attendance. These programs will be included in the systematic review but analyzed separately. Programs that target teenage parents are distinct in their focus on a particular target population, but also involve program components (e.g., daycare services) that are not offered in dropout programs targeted toward general population students. In addition, policymakers would not typically choose between a teen parent program and a general program when making programming decisions. Rather, selections would be made from among a group of teen parent programs for one target population, and from among a group of general programs for another target population.

1.3 PREVIOUS REVIEWS ON DROPOUT INTERVENTIONS

There have been a handful of systematic reviews on the effects of prevention and intervention programs on school dropout and completion outcomes. However, the restrictive inclusion criteria and methodological weaknesses of these reviews preclude any confident conclusions about the effectiveness of the broad range of programs with dropout outcomes, or the potential variation of effectiveness for different program types or participant populations. For instance, the U.S. Department of Education's What Works Clearinghouse report on dropout prevention found only 15 qualifying studies that reported outcomes on direct measures of staying in school or completing school (http://ies.ed.gov/ncee/wwc/reports/dropout/topic/#top). This report, however, restricted discussion to interventions in the United States and did not include a meta-analysis of program effectiveness or examine potential moderators of program effectiveness. Another review on best practices in dropout prevention summarized the results of 58 studies of dropout programs (ICF International, 2008). That report presented effect sizes primarily for individual program types and, like the What Works Clearinghouse report, did not examine potential moderators or examine the influence of study method on effect size. The report also presented a narrative review of important variables associated with implementation quality, but implementation quality was not analyzed in a meta-analysis framework.

Although not a systematic review, the National Dropout Prevention Center/Network produced a detailed report on exemplary and evidence-based programs that address risk factors for school dropout (Hammond, Linton, Smink, & Drew, 2007). A strength of the report was that it categorized programs into primary, selected, and indicated interventions and described the different risk factors targeted by the program strategies. This information can be useful for practitioners who wish to select programs to suit the characteristics of their particular student population or setting. While the report covered a wide range of program strategies, the authors did

not report effect sizes or perform a meta-analysis, making it difficult to judge the effectiveness of the interventions or compare the effectiveness across programs.

Two other systematic reviews have focused on the effectiveness of prevention and intervention programs to reduce school dropout or increase school completion (Klima, Miller, & Nunlist, 2009; Lehr et al., 2003). In their review, Lehr et al. (2003) identified 17 experimental or quasi-experimental studies with school enrollment status outcomes. This review was completed seven years ago, and thus does not include the most recent studies. The authors did not perform a meta-analysis because they felt that the dependent variables differed too greatly across studies to create meaningful aggregates. This circumstance obviously prevented the authors from examining the differential effectiveness of programs with different program or participant characteristics. In a more recent review, Klima et al. (2009) identified 22 experimental or quasi-experimental studies with dropout, achievement, and truancy outcomes. However, this review excluded programs for general "at-risk" populations of students (e.g., minority or low socioeconomic status samples), as well as programs with general character-building, social-emotional learning, or delinquency/behavioral improvement components. These exclusion criteria therefore limited the conclusions that could be drawn about the broader range of programs that aim to influence school dropout and completion outcomes. Further, this review only presented mean effect sizes for different types of interventions, and did not examine the potential variation of effects for different subject populations.

Nonetheless, the findings of the Klima et al., (2009) and Lehr et al., (2003) reviews have some similarities. Both teams highlight the dearth of high-quality research on dropout programs, and mention especially the lack of key outcomes such as enrollment (or presence) at school and dropout. Both reviews demonstrate that some of the included programs had positive effects on the students involved. Lehr and her colleagues did not identify specific programs that were particularly effective or ineffective, but focused rather on implementation integrity as a key variable and emphasized the importance of strong methodologies for future research on dropout programs. Klima and colleagues concluded that the programs they reviewed had overall positive effects on dropout, achievement, and attendance/enrollment. They highlighted school restructuring programs, such as schools-within-schools, as particularly effective. The Klima review also suggests that alternative school programs, that is, programs in separate school facilities, were ineffective. Overall, these two reviews identify several important potential moderators that will be included in the present review. These include implementation quality, type of program, and whether programs are housed in typical school facilities or in alternative school locations.

1.4 OBJECTIVES OF THE REVIEW

The objective of our systematic review is to summarize the available evidence on the effects of prevention and intervention programs aimed at primary and secondary students for increasing school completion or reducing school dropout. The primary focus of the analysis will be the comparative effectiveness of different programs and program approaches in an effort to identify those that have the largest and most reliable effects on the respective school participation outcomes, especially with regard to differences associated with program type, implementation quality, and program location or setting. In addition, evidence of differential effects for students with different characteristics will be explored, e.g., in relation to age or grade, gender, and race/ethnicity. Because of large ethnic and socioeconomic differences in graduation rates, it will be particularly important to identify programs that may be more or less effective for disadvantaged students. A final objective of the review is to summarize the evidence on the effects of interventions aimed at increasing school completion for pregnant and parenting teens.

The ultimate objective of this systematic review is to provide school administrators and policymakers with an integrative summary of research evidence that is useful for guiding programmatic efforts to reduce school dropout and increase school completion for all types of students.

2 Methods

2.1 TITLE REGISTRATION AND REVIEW PROTOCOL

The title for this systematic review was published in The Campbell Collaboration Library of Systematic Reviews on January 18, 2010. The review protocol was published on November 9, 2010. Both the title registration and protocol are available at: http://www.campbellcollaboration.org/library.php.

2.2 ELIGIBILITY CRITERIA

To be eligible for inclusion in the review, studies were required to use an eligible intervention directed toward an eligible participant sample, report data that permitted calculation of a numeric effect size for at least one eligible outcome variable, and employ an eligible research design, as outlined below. Note that the eligibility criteria and coding scheme (described below) were identical for the general population and teen parent programs.

2.2.1 Interventions

Eligible interventions included school-based or school-affiliated psychological, educational, or behavioral prevention or intervention programs, broadly defined, that involved actions performed with the expectation that they would have beneficial effects on student recipients. School-based programs are defined as those that were administered under the auspices of school authorities and delivered during school hours. School-affiliated programs are those that were delivered with the collaboration of school authorities, possibly by other agents, e.g., community service providers, and which may have taken place before or after school hours and/or off the school grounds. Community-based programs that were explicitly billed as dropout prevention or intervention programs were also eligible whether or not a school affiliation was evident. Other community-based programs that included dropout among their goals or intended outcomes, but for which dropout or related variables were not the main focus, and which had no evident school affiliation were not eligible. Programs that were solely medical or pharmacological in nature were not eligible. Broad programs and policies that were implemented at the district level where no intervention could be identified as occurring at the school level, such as district catchment area restructuring, were not eligible.

2.2.2 Participant Populations

Eligible participant populations included school-aged youth, defined as those expected to attend pre-k to 12th grade primary and secondary schools, or the equivalent in countries with a different grade structure, corresponding to approximately ages 4-18. The age or school participation of the sample was required to be presented in sufficient detail to allow reasonable inference that studies met this requirement. Studies with participant populations including recent dropouts between the ages of 18-22 were also eligible if the program under study was explicitly oriented toward secondary school completion or the equivalent.

General population samples of school-age children were eligible, as were samples from populations broadly at risk because of economic disadvantage, individual risk variables, and closely related risk factors (e.g., inner city schools, students from low socioeconomic status families, teen parents, students with poor attendance records, students who had low test scores or who were over-age for their grade). Samples consisting exclusively of specialized populations, such as students with mental disabilities or other special needs, were not eligible. However, inclusion of some such individuals in a broader sample in which they were a minority proportion did not make that broader sample ineligible. Students with learning disabilities, such as dyslexia, that generally do not require them to be in specialized schools or classrooms (i.e., they attend mainstream classes and typical schools) were considered eligible.

2.2.3 Research Designs/Comparison Conditions

Eligible studies were required to use an experimental or quasi-experimental design with at least 10 subjects in each of the intervention and control groups at the time of assignment. The following research designs were eligible:

a. Participants were randomly assigned to intervention and control conditions or assigned by a procedure plausibly equivalent to randomization.
b. Participants in the intervention and control conditions were matched and the matching variables included a pretest for at least one qualifying outcome variable. However, if the qualifying outcome variable did not lend itself to meaningful pretest or the pretest values could be assumed zero (such as dropout when all students were enrolled in school at baseline), but the groups were matched on other variables plausibly related to risk for dropout, the study was still eligible. For this purpose, use of pretest or initial risk variables as statistical controls, e.g., in an ANCOVA or multiple regression analysis, was considered the equivalent of matching.
c. If participants were not randomly assigned or matched, the study must have reported both a pretest or relevant baseline risk variables and a posttest on at least one qualifying outcome variable with sufficient statistical information to derive an effect size or to estimate group equivalence from statements of statistical significance, or provide evidence of equivalence on key risk variables and/or student characteristics.

Studies that employed designs in which more than one intervention group was compared to a single control group were eligible; in these cases, effect sizes were calculated for each intervention group compared to the control group.

Posttest only non-equivalent comparisons (not randomized, matched, or demonstrating equivalence) were not eligible for inclusion in the review. Single-group pretest-posttest designs were not eligible for inclusion in the review.

2.2.4 Outcomes

All eligible studies were required to assess intervention effects on at least one outcome variable that represented school completion or dropout, or was a close proxy measure or recognized precursor for dropout. Qualifying outcome variables were those that fell in or were substantially similar to the following categories:

- School dropout;
- Secondary school completion, high school graduation, or attainment of GED or similar certificate[1];
- Absences or truancy;
- Enrollment/non-enrollment in school.

Studies in which the majority of children were under middle school age (approximately 5th grade in the United States or age 11) were required to have either a school completion or dropout outcome, or have attendance measures that were assessed in middle or high school. This criterion had the intended effect of excluding the large number of school-based studies for elementary age students that focus mainly on attendance but have no specific dropout focus or dropout outcome.

2.2.5 Other Study Characteristics

Eligible studies were required to be relatively modern in order for the research to be applicable to contemporary students. Therefore, the date of publication or reporting of the study was required to be 1985 or later even though the research itself may have been conducted prior to 1985. If, however, there was evidence in the report that the research was actually conducted prior to 1980 (more than five years before the 1985 cutoff date), then the study was not considered eligible for inclusion.

Eligible studies were required to report sufficient quantitative data to compute an effect size on an eligible outcome. In addition, the variables involved in the effect size must have had a known direction of scoring, i.e., whether high or low values represented favorable or less favorable results.

Studies could have been published or reported in any language and conducted in any country as long as they met all other eligibility criteria.

[1] General Educational Development (or GED) tests are a group of five subject tests which, when passed, certify that the taker has American or Canadian high school-level academic skills. The GED is also referred to as a General Education Diploma, General Equivalency Diploma, or Graduate Equivalency Degree.

2.3 LITERATURE SEARCH

2.3.1 Information Sources

An attempt was made to identify and retrieve the entire population of published and unpublished studies that met the inclusion criteria summarized above. The following electronic bibliographic databases were searched:

1. Dissertation Abstracts International searched 1985 to January 2010
2. Education Abstracts searched 1985 to January 2010
3. PsycInfo/ERIC searched 1985 to January 2010
4. SSCI searched 1985 to January 2010
5. DissOnline searched 1985 to March 2011
6. Index to Theses searched 1985 to March 2011
7. Theses Canada searched 1985 to March 2011
8. Australian Education Index searched 1985 to April 2011
9. Canadian Education Index (Canadian Business and Current Affairs Education) searched 1985 to April 2011
10. British Education Index searched 1985 to April 2011

The search terms were adapted to each database, and varied in complexity depending on database capabilities. Generally, two blocks of search terms were used. One block included the set of overlapping terms describing the population, the intervention and/or the outcomes of interest. The second block described the research designs. When possible, a series of exclusion terms was also used. Because the term "dropout" is used commonly in other research domains (e.g., as in dropout from mental health therapy), the list of exclusion terms allowed us to exclude large numbers of false positives. The search terms used for PsycINFO and ERIC within the CSA database were as follows, limited to the date range of 1985-2010 and using CSA age limits from childhood through young adulthood:

((DE=((school dropouts) or (school attendance) or (truancy) or (school graduation) or (high school graduates))) or (KW=((school complet*) or (GED) or (general education development) or (high school diploma) or (dropout*) or (alternative high school*) or (drop* out) or (career academ*) or (school NEAR absen*) or (chronic* NEAR absen*) or (school enrollment) or (high school equivalency) or (school failure) or (high school reform) or (educational attainment) or (grade promotion) or (grade retention) or (school nonattendance) or (graduation rate) or (school refusal))))

and

(KW=((intervention) or (quantitative) or (program evaluation) or (random*) or (prevent*) or (pilot project*) or (youth program*) or (counseling) or (guidance program*) or (summative evaluation) or (RCT) or (clinical trial) or (quasi-

experiment*) or (treatment outcome*) or (program effect*) or (treatment effect*) or (evaluation) or (experiment*) or (social program) or (effective*)))

and not

((KW=((post-secondary) or (undergraduate) or (doctoral) or (inmate) or (schizophrenia) or (traumatic brain injury) or (autis*) or (abuse) or (antidepressant*) or (unipolar depression) or (risperidone) or (chronic illness) or (major depressive disorder) or (bulimia) or (buprenorphine) or (malaria) or (heroin) or (cancer) or (major depression) or (Massage therapy) or (fibromyalgia) or (Paroxetine) or (clomipramine) or (olanzapine) or (tuberculosis) or (spinal cord injury) or (epilep*) or (antiepileptic) or (HIV) or (psychosis) or (OCD) or (obsessive-compulsive) or (EEG) or (PTSD) or (tourette*) or (insomnia) or (obes*) or (anorexia) or (methadone) or (borderline personality disorder) or (mental retardation))) or (DE=((higher education) or (college students) or (treatment dropouts) or (employee absenteeism))))

In addition to the aforementioned electronic bibliographic databases, we also searched numerous other research registers, organization websites, and other online libraries in an attempt to locate grey or unpublished literature and any other narrative or systematic reviews on the topic. These other information sources included:

1. Campbell Collaboration Library searched through January 2010
2. Cochrane Collaboration CENTRAL Library searched through January 2010
3. OpenSIGLE Library searched through February 2010
4. NTIS Library searched through February 2010
5. National Dropout Prevention Center/Network Website searched through March 2010
6. National Institutes for Health Research, National Research Register searched through March 2010
7. Canadian Evaluation Society Grey Literature Database searched through March 2010
8. NLM Gateway searched through March 2010
9. Social Care Online Database searched through March 2010
10. American Evaluation Association Library searched through March 2010
11. CERUK Plus searched through March 2011
12. EPPI-Centre Systematic Reviews searched through March 2011

We also checked the bibliographies of all screened and eligible studies for any potentially relevant reports not identified in the electronic searches, as well as the bibliographies in prior narrative reviews and meta-analyses of school dropout programs. Continued contact was maintained with the expert consultant, Sandra L. Christenson from the University of Minnesota, regarding potentially eligible studies that had been previously reported or were currently in progress.

2.3.2 Study Selection

Abstracts and/or titles were retrieved for all search hits from the bibliographic databases. Staff at the Peabody Research Institute (including authors Wilson, Tanner-Smith, Steinka-Fry, and Morrison) screened all abstracts and/or titles to first eliminate any clearly irrelevant study reports (e.g., theoretical articles, single group studies, non-dropout related articles). Full text versions of reports were retrieved for all remaining reports that were not clearly or explicitly ineligible. If there was any ambiguity about the potential eligibility of a report based on the abstract (or title, when abstracts were not available), we erred on the side of inclusivity and retrieved the full text report to make the eligibility decision. The authors then screened the full text reports to make final eligibility decisions for each report, following the eligibility criteria outlined above.

2.3.3 Data Collection Process

All coding was conducted by six coders (including authors Wilson, Tanner-Smith, Steinka-Fry, and Morrison) who participated in several weeks of initial training led by Wilson as well as ongoing coding meetings throughout the course of the project. During the initial training, all coders coded at least two of the same studies and then convened as a group to resolve any coding discrepancies until 100% reliability was attained on all coded variables. After the training period, all coding questions were addressed in weekly team meetings and decided upon with the group. Wilson also double-coded all program types and resolved any coding discrepancies with the individual coder. In addition, Wilson, Tanner-Smith, Steinka-Fry, and Morrison double-coded a random subset of each others' studies, and resolved any discrepancies with the individual coder. If the two coders could not resolve a discrepancy, they raised the issue in the weekly team meeting to be resolved with the entire group.

Coders entered data directly into a FileMaker Pro database using computer screens tailored to the coding items and with help links to relevant sections of the coding manual. Effect size calculations were built into the data entry screens for the most common statistical representations, and David Wilson's online effect size calculator (http://www.campbellcollaboration.org/resources/effect_size_input.php) and Shadish, Robinson, & Lu's *ES* calculator program were used to estimate effect sizes from less common representations.

2.3.4 Coding Protocol

Eligible studies were coded on numerous variables related to study methods, the nature of the intervention and its dosage and implementation quality, the characteristics of the participant samples, the outcome variables and statistical findings, and contextual features such as setting, year of publication, and the like. A detailed coding manual is included in the supplementary materials.

2.4 ANALYTIC METHODS

2.4.1 Effect Size Indices

Odds ratios were used as the effect size metric for dropout and other binary outcome measures (e.g., graduation, school enrollment). All effect sizes were coded such that larger effect sizes represent positive outcomes (e.g., less school dropout, higher school enrollment). All computations described in the results below that used odds ratios were carried out with the natural logarithm of the odds ratio.

The meta-analyses reported in this review only summarize results for school dropout and school completion outcomes, all of which were dichotomous outcome measures that could be represented with odds ratios. Many of the secondary outcomes that were coded from the reports, but not analyzed in the current review, were measured on a continuous scale (e.g., absences, attendance, academic performance). These outcomes were measured using standardized mean difference effect sizes corrected for small sample bias (Hedges, 1981). However, none of those results are reported in the present review.

2.4.2 Synthesis Methods

All analyses were inverse variance weighted using random effects statistical models that incorporate both the sampling variance and between studies variance components into the study level weights. Random effects weighted mean effect sizes were calculated for all studies using 95% confidence intervals. Estimates of Cochrane's Q and τ^2 were used to assess heterogeneity in the effect sizes.

Because many studies reported multiple effect sizes on the same participant samples (e.g., two measures of school dropout were reported), it was not feasible to assume independence of the effect size estimates – an assumption used in most standard meta-analysis methods. Therefore, all meta-analyses reported in this review used robust variance estimation techniques that account for these statistical dependencies, as outlined recently in Hedges, Tipton, & Johnson (2010).

2.4.2.1 *Outliers*

The distribution of logged odds ratio effect sizes was examined for outliers. Outliers were defined as values that fell more than three interquartile ranges (IQR) above the 75th percentile or below the 25th percentile of the distribution (Tukey, 1977). Two outliers were identified and were Winsorized to the next closest value.

2.4.2.2 *Cluster Adjustments*

Thirty-five of the 504 effect sizes analyzed below required cluster adjustments because assignment to treatment conditions was performed at the school level and data were reported at the individual level. Because none of these studies provided

between-school intra-class correlations, an ICC for high school graduation of .24 from Altonji and Mansfield (2010) was used to conduct cluster corrections of the sample sizes per Hedges (2007). The standard errors for these odds ratios were then recomputed using the new effective sample sizes.

2.4.2.3 *Missing Data*

A small number of studies were missing data on the method, participant, or program variables used in the final analyses; missing values were imputed using an expectation-maximization (EM) algorithm in SPSS (Graham, Cumsille, & Elek-Fisk, 2003). Sensitivity analyses were performed for all analyses using imputed missing values.

2.4.3 Moderator Analyses

The main objective of the review was to describe the direction and magnitude of the effects of different interventions on the outcomes in a manner that allows comparative effectiveness to be assessed. Therefore, we conducted moderator analyses using meta-regression models to identify the characteristics of the interventions and participants associated with the largest and smallest effects. As outlined in the original review protocol, the a priori effect size moderators of interest were: program modality/type, implementation quality, program duration, program location, gender and ethnic mix of participants, and age or grade-level of participants.

2.4.4 Publication Bias Analyses

Funnel plots and regression based tests (Peters et al., 2006), were used to assess the possibility of publication bias in the analytic sample and its potential impact on the findings of the review.

3 Results

3.1 LITERATURE SEARCH & STUDY IDENTIFICATION

Figure 7.1 in Appendix II shows a flow diagram of all identified research reports for the current review. A total of 23,677 study reports were identified during the search phase: 20,246 were identified via searches of electronic databases and 3,431 were identified from other sources (citation searching, etc.). Abstracts (or titles, when abstracts were not available) of all reports were screened for initial relevance. Of those, 20,767 were excluded as being clearly irrelevant or ineligible for inclusion in the review. The large number of exclusions at this stage resulted from the search terms used in the electronic database searches that aimed to be as liberal as possible in order to locate all relevant reports.

We attempted to obtain full text copies for all other reports—including those that may have been plausibly eligible based on the abstract or title, and those for which it was unclear whether they were eligible or not. At present, a total of 41 full-text reports have been deemed un-retrievable based on repeated attempts to locate the reports through the Vanderbilt University interlibrary loan service and email contact with report authors. The remaining 75 un-retrieved reports have been requested through the Vanderbilt University interlibrary loan service and have not been located at the time of this writing. We believe that most (40) of these reports may be linked with studies that are already coded or eligible in our database (i.e., studies with multiple reports describing their results), and thus, if located, are unlikely to provide new information beyond what we have already extracted from the other retrieved reports for those studies. If any of the remaining unretrieved reports turn out to be eligible, their results will be included in the next update of this review.

A total of 2,794 full text reports were retrieved and screened for eligibility. Among those, 1,673 were not eligible, 230 were review articles whose reference lists were harvested for potentially relevant citations, and 891 were deemed eligible. The current review presents results for the 167 coded studies based on the 548 eligible study reports with school dropout outcomes. Table 6.1 in Appendix I provides additional detail on the number of reports deemed not eligible, by specific inclusion criteria. A bibliography of all eligible studies is included in the supplementary materials.

The published protocol for this review proposed to analyze both school dropout and school attendance outcomes in the same systematic review, and indeed the eligibility criteria outlined above list school attendance as an eligible outcome variable. Several of the previous systematic reviews we read when preparing the protocol complained about the dearth of high quality research on dropout programs and included only a small number of studies in their reviews. We included attendance outcomes in our original analysis plan based on the expectation that the literature would yield fewer than 100 eligible studies. However, after we conducted our literature search, it became clear that there was a much larger sample of eligible studies available, most of which were unpublished technical reports. Given this circumstance of locating many more eligible studies than we expected, the current review focuses solely on the 167 studies reporting the primary outcome of interest—school dropout or school completion. The remaining 303 reports (239 studies) that reported school attendance outcomes only will be the topic of a second systematic review. We have further identified 40 reports (22 studies) that were conducted in developing countries (aimed at increasing school enrollment and attendance through school feeding programs or conditional cash transfer programs); a third review is planned for the developing country studies.

3.2 SELECTING STUDIES FOR ANALYSIS

Of the 167 studies included in the review, 152 involved programs geared toward general or at-risk populations of school-aged youth, while 15 focused exclusively on teen parents.[2] Given the distinct focus and target population for the teen parent programs, all analyses were conducted separately for the general programs and the teen parent programs.

Furthermore, many of the 167 studies included in the review reported results separately for different research sites and/or for different demographic subgroups of participants. When we could be sure that the different research sites or subgroups contained no overlapping subjects, we treated them as independent samples and calculated effect sizes separately for the independent samples. For the general population programs, the 152 studies produced 317 independent samples. For the teen parent programs, the 15 studies produced 51 independent samples.

3.3 GENERAL DROPOUT PROGRAMS

3.3.1 Types of Dropout Programs

Table 3.1 provides descriptions of the most common types of general dropout programs, listed in order of prevalence. Note that while a focal program category

[2] Some of the teen parent programs were open to young fathers as well as young mothers; however, none of the individuals in the study samples for the teen parent programs were male.

was selected for each program in the review, most programs implemented multiple components and many included components from several of the categories shown in the table. School or class restructuring programs were the most common in the review. The simplest form of these programs involved creating smaller classes or lower student-teacher ratios. The more complex programs in this category involved larger scale school restructuring activities that could include some combination of smaller classes, blocked schedules, personalized learning settings in which students spend more time with fewer teachers, and/or "academies" focused around grade level (e.g., 9th grade academies) or career foci.

Table 3.1: Program Types–General Dropout Programs

Program Type	Description	N of samples (*k*)	N of effect sizes (*n*)
School or class restructuring	Small learning communities, block schedules, career academies, small class size	91	105
Vocational training	Coursework, internships, or employment oriented toward work or career interests	49	51
Supplemental academic services	Remedial education, tutoring, homework assistance, etc.	28	32
Community service	Programs involved planning and carrying out a community service project (commonly coupled with a weekly life skills curriculum)	24	24
Mentoring, counseling	Programs provided adult mentors or trained counselors for students. Though mentors focused more on career/work, both mentors and counselors dealt with students' personal issues	23	27
Alternative schools	Schools designed to provide educational and other (e.g., behavioral) services to students whose needs aren't adequately addressed in traditional schools. Typically for students pushed out of regular schools	22	30
Attendance monitoring & contingencies	Monitoring and services to increase attendance; some offer financial incentives	20	26
College-oriented programming	College preparatory curriculum, college-oriented academic advising	18	25
Multi-service package	Large, comprehensive programs; often included academic, vocational, & case management	12	23
Skills training, including CBT	Generally oriented toward improving self-esteem or attitudes about school, or preventing drug use	12	12
Case management	Programs revolved around connecting students and families with appropriate services	10	17
Other	Recreational, residential services for homeless, etc.	8	13

The next most prevalent type of program focused on vocational training, work-related coursework, career exploration, internships, and/or paid employment for students. Supplemental academic services (e.g., tutoring, homework assistance) and community service programs were also common. Another group of programs involved mentoring and counseling. These programs often included mentoring or counseling with regard to career planning, and many including personal counseling as well. Alternative schools were relatively common in the literature (with 22 studies). Alternative schools are generally separate educational facilities which provide educational services to students who have been excluded from traditional school settings. Most students in alternative school settings are at risk for dropout. There were several attendance monitoring and behavioral contingency programs that applied rewards and punishments based on attendance, school performance, school attendance, or other behaviors. The remaining program types included college-oriented programming, case management, counseling, skills training, and multi-service packages that offered a broad range of diverse services.

3.3.2 Methodological Characteristics

Table 3.2 presents the methodological, participant, and program characteristics for the general dropout programs. In terms of methodological characteristics and quality, it is noteworthy that a large proportion of the literature on dropout programs is not published in peer-reviewed journals, but rather reported in technical reports (55%) or dissertations (22%). Many (if not most) of the technical reports were of high quality and published by government and non-government agencies conducting large scale evaluations of dropout programs. Indeed, a greater proportion of the technical reports used random assignment designs (49%) than did the journal articles (24%). Attrition averaged about 11% across the studies of general programs that reported it. About 25% of the effect sizes extracted from the general programs studies were calculated using adjusted data. That is, the authors presented results that were adjusted for differences between intervention and comparison groups at the outset of the study. Finally, the roles of the evaluators in the research studies varied. Most common were independent evaluators, but some evaluators were more closely involved in the program delivery.

3.3.3 Program Dosage, Format, and Location

School locations were the most common settings for dropout programs, as would be expected. Over half of the programs were delivered in classroom settings during regular class time. In addition, 32 studies involved programs delivered at school, but not in typical classrooms (e.g., programs in which students were pulled out of class for special services). Another 40 programs were delivered after school. Thirty-four programs were delivered in community settings. The community-based programs all targeted dropout directly and were more likely to include case management or mentoring. Finally, 36 studies were delivered in multiple settings, typically both community and school-based components were present.

Table 3.2: Characteristics of Studies of General Dropout Programs

Study Characteristics					
General Characteristics	*mean*	*sd*	Program Characteristics		
Publication year	1994	7	*Role of Evaluator*	*k*	*%*
			Delivered program	6	2
			Involved in planning, not		
Form of Publication	*k*	*%*	delivery	27	9
Journal article	70	22	Influential, but no direct role	33	10
Technical report	173	55	Independent evaluator	250	79
Dissertation	70	22	*not reported in 1 sample*		
Other	4	2			
			Program Site	*k*	*%*
Country of Study	*k*	*%*	School classroom	175	55
United States	309	97	School, but not classroom	32	10
UK, Canada	8	3	After school	40	13
			Community site	34	11
Method Characteristics			Mixed or multiple sites	36	11
Assignment Method	*k*	*%*			
Random (individual)	123	39	*Implementation Quality*	*k*	*%*
Random (group)	3	<1	Clear problems	81	26
Matched	118	38	Possible problems	65	21
Nonrandom, nonmatched	73	23	No apparent problems	171	54
	mean	*sd*	*Program Dosage*	*mean*	*sd*
Attrition	11%	20%	Duration (weeks)	91	73
not reported in 53 samples			Hours per week	17	15
			Total contact hours	1452	1733
*Adjusted Data Used in OR**	*n*	*%*	*not reported in 16 samples*		
Yes	96	25			
No	287	75	*Frequency of Program*	*k*	*%*
			Less than weekly	16	5
Participant Characteristics	*mean*	*sd*	Once a week	31	10
Percent male	50	23	Twice a week	33	11

not reported in 29 samples			3-4 times a week	21	7
	mean	*sd*	Daily (5 days/week)	191	64
Percent White	33	32	Continuous	5	2
Percent Black	39	34	*not reported in 20 samples*		
Percent Hispanic	22	28			
Percent other minority	9	19	Dependent Variables*	*n*	*%*
not reported in 49 studies			Dropout	238	62
	mean	*sd*	Graduation	84	22
Average age	15	3	Graduation or GED	34	9
Average grade level (US)	9	2	Enrolled (not dropped, graduated)	29	8

Notes: Number of effect sizes (n = 385); Number of independent samples (k = 317)
**Reported at effect size level (n = 385)*

Treatment dosage was fairly intensive for the dropout programs. Duration averaged 91 weeks (or about 2 school years) and about 64% of programs involved daily contact with students (i.e., 5 days/week). However, many of the studies included in the review (47%) mentioned some problems with implementation. There was considerable variability across programs in what (and how much) they reported about program implementation. For instance, some studies simply stated that programs experienced problems with implementation, but provided no additional information on specific types of problems. More commonly, studies reported difficulties related to staffing or funding, problems with administrator buy-in, or other structural difficulties in implementing the program (e.g., computer access, lab space). Many studies also reported more serious implementation problems related to control group contamination effects, intervention group participants receiving significantly less intervention contact than intended (e.g., 39% completed fewer than 75% of planned sessions), or intervention group participants not receiving entire intervention components (e.g., planned child care, case management, or mentoring services were never provided to participants).

3.3.4 Participant Characteristics

The intervention and comparison groups were typically about half male, as would be expected for a group of primarily school-based studies. Mixed ethnicity groups were common, with minority youth somewhat over-represented in relation to their proportions in the general population in the United States. While it is not shown in the table, we coded text descriptions of the risk of dropout for each study. Many of the studies focused on at-risk groups, i.e., students with poor performance, poor attendance, and the like. That minority youth in the United States are more likely to be disadvantaged and experience other risk factors like poor academic performance may explain their overrepresentation in the research studies here. The average age of the study participants was about 15 and the average grade level was 9th grade (typically, the first year of four-year high school in the United States). The age range was considerable, with a few programs targeting elementary age students but most focusing on the high school age range.

3.3.5 Synthesis of Results for General Dropout Programs

As mentioned above, the independent study samples frequently reported results on multiple outcomes and for multiple waves of data collection. For the analyses presented here, we allowed for multiple dependent variables from the same study sample, as long as they represented school dropout, school completion, or school enrollment but selected only a single wave of measurement if multiple waves were available. We selected the first wave of measurement after program completion, or the wave closest to the end of the program if no post-program waves were given. The dependencies resulting from the use of multiple dependent variables from the same study sample are handled using robust variance estimation techniques (Hedges, Tipton, & Johnson, 2010).

The 317 general programs provided 385 odds ratios for analysis. High school dropout was the most common dependent variable (n = 238 odd ratios), with graduation outcomes representing 22% (n = 84) of the odds ratios. Thirty-four odds ratios indexed graduation or GED attainment, and 29 odds ratios indexed school enrollment. School enrollment was defined as not dropped out or graduated and still attending school. The random effects weighted mean odds ratio for the general programs was 1.72 (*95% CI* 1.56, 1.90). The mean odds ratio was positive and statistically significant, indicating that participants in dropout programs exhibited lower dropout rates and higher graduation rates than participants in the comparison conditions (see Table 3.3). To make interpretation more straightforward, we can convert the mean odds ratio into percentage terms using an estimate of the proportion of dropouts in the comparison groups. Across all comparison groups in the general programs in the review, the average percentage of students who dropped out was 21%. Using 21% as our baseline, the mean odds ratio of 1.72 translates into a dropout rate for program students of 13%. That is, while 21% of students in regular educational programs dropped out of high school, only 13% of students participating in dropout programs dropped out of high school, an 8 percentage point difference.

Table 3.3: Mean Odds Ratio Effect Sizes and Heterogeneity Statistics for General Programs

	Mean OR	*95% CI*	*n*	*k*	Q_E	τ^2
General Dropout Programs	1.72	1.56, 1.90	385	317	3065	.56

Notes: Mean effect size and confidence interval estimated with robust standard errors to account for dependent effect sizes estimates. n– number of effect sizes; k – number of unique samples; Q_E – weighted sum of squares heterogeneity statistic that accounts for dependent effect size estimates, τ^2 – method of moments between studies variance component that accounts for dependent effect size estimates.

The Q_E is the heterogeneity statistic for the distribution of odds ratios that accounts for the dependent effect size estimates. For the general programs Q_E = 3065 (df = 316-384, $p < .05$) indicating significant heterogeneity in the distribution of odds ratios. The results across the 317 independent study samples exhibited more variability than can be explained from sampling error alone. Because some studies had larger or smaller treatment effects than others, our next step was to examine the relationships of our odds ratios to the various study characteristics reported in Table 3.2 above.

3.3.6 Moderator Analysis for the General Dropout Programs

Moderator analysis was performed using meta-regression with robust standard error estimation. The robust standard errors used an estimate of a between-effect within-study correlation, ρ, estimated at .80; sensitivity analyses indicated the findings were robust across different reasonable estimates of ρ.

Three meta-regression models were run for the general dropout programs. Model I included characteristics of the study methods. The methodological characteristics, as we shall see below, were reliably associated with study results. While perhaps not

substantively interesting, it is important to include these variables in moderator analyses to get a better sense of the true effects of the programs, net of the influence of study method. We included nine methodological variables in the meta-regression models: two dummy codes representing random assignment and matched designs (with the reference group including non-random, non-matched designs), percent attrition, the mean pretest group equivalence odds ratio from the study sample, a dummy code for whether the data used to calculate the effect size were adjusted for covariates, three dummy codes for the dependent variable operationalization (with the reference category representing the school enrollment outcomes), and a four-level ordinal variable representing the degree of independence of the evaluator (higher scores indicate more independence).

In Model II, three additional subject variables were included: percent male, percent white, and average age. While participant characteristics are important from a policy perspective, inasmuch as dropout programs may be more or less effective for students with different characteristics, our intent here is to examine in more detail the characteristics of the programs that may be more alterable, such as implementation, dosage, and program components.

In Model III, dosage, format, and implementation variables were added. These included implementation quality, rated on a 1-3 scale with higher scores indicating no apparent implementation problems; and, three dummy codes for program format, classroom-based, school-based but not in a classroom (either pull-out or after school), and mixed or multiple formats. The reference group included programs delivered in community settings such as clinics and community centers. Two variables represented program dosage: total program length in weeks and the number of service hours per week. Note that many of the moderator variables we included in the three models were correlated with each other and with the different program types (the correlation matrix is presented as Table 6.2 in Appendix I).

Several interesting findings are evident from the meta-regression models shown in Table 3.4. Random assignment and matched designs tended to produce smaller effect sizes than the non-random, non-matched designs in the reference group, even while controlling for the initial group equivalence and covariate adjustments. The random assignment designs produced significantly smaller effect sizes in the first two models. When the dosage and implementation variables were included, however, the random assignment dummy variable was no longer significant. Attrition and the different operationalizations of the outcome variables (dropout, graduation, etc.) were not generally associated with the odds ratios, indicating that the odds ratios were generally similar regardless of the amount of attrition or the particular form of the outcome variable. The initial equivalence of the two groups was significantly associated with the outcomes in Model III. Larger pretest differences between the intervention and control groups were associated with larger posttest differences between the groups, regardless of the method of assignment and the other variables in the model. Controlling for these initial differences over and

above the assignment method, therefore, is important to better isolate the effects of the dropout programs, net of initial nonequivalence between the two groups. Odds ratios that were computed using data adjusted for covariates (yet another way of creating more similar intervention and comparison groups) were also significantly smaller than those that were unadjusted.

Table 3.4: Meta-Regression Results Predicting Logged Odds Ratio Effect Sizes – General Programs

	Model I			Model II			Model III		
	B	se	95% CI	b	Se	95% CI	b	se	95% CI
Methodological Characteristics									
Random assignment (1=yes)	-.29*	.15	-.57, -.00	-.29†	.15	-.57, .00	-.03	.15	-.33, .27
Matched groups design (1=yes)	-.19	.16	-.50, .12	-.18	.16	-.50, .13	-.01	.15	-.31, .29
Attrition (%)	-.04	.24	-.52, .44	-.02	.24	-.50, .46	.07	.25	-.42, .56
Mean group equivalence OR	.15	.16	-.16, .45	.22	.19	-.15, .59	.42*	.19	.03, .80
Adjusted data for OR calc (1=yes)	-.36*	.09	-.55, -.18	-.32*	.09	-.50, -.13	-.39*	.11	-.60, -.18
Dropout outcome (1=yes)	-.06	.16	-.37, .26	-.06	.17	-.38, .27	-.08	.15	-.37, .21
Graduation outcome (1=yes)	.13	.19	-.24, .50	.18	.20	-.20, .57	.15	.19	-.23, .52
Graduated/GED (1=yes)	.21	.20	-.19, .61	.13	.20	-.27, .53	.18	.21	-.24, .59
Independent evaluator (1-4)	-.12	.09	-.29, .05	-.15†	.09	-.32, .03	-.12	.08	-.28, .05
Participant Characteristics									
Percent male				.01	.22	-.41, .44	.07	.20	-.33, .47
Percent White				.03	.17	-.31, .37	.05	.17	-.28, .37
Average age				.03	.02	-.02, .08	.04†	.02	-.01, .09
Program Characteristics									
Implementation quality (1-3)							.30*	.06	.18, .41
Classroom program (1=yes)							.45*	.12	.20, .69
School, not classroom (1=yes)							.23	.14	-.05, .51
Mixed program sites (1=yes)							.29†	.16	-.02, .60
Program duration (in weeks)							.00	.00	.00, .00
Hours of service per week							-.01†	.00	-.01, .00

Notes: Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates. b – unstandardized regression coefficient.

† $p < .10$. * $p < .05$.

In terms of participant characteristics, there were few differences in treatment effects associated with different participant characteristics. That is, programs with more males were not any more or less successful than programs with more females. The same was true for studies with different proportions of white students. Majority and minority students seemed to benefit about equally from dropout programs. Age was moderately associated with treatment effects in the third model (b = .04; p < .10), with programs with older students tending to show better outcomes than those with younger students.

The format, dosage, and implementation variables also provide some interesting results. However, the different types of dropout programs, e.g., school restructuring or vocational training, tend to have different dosage levels as part of their overall treatment package. Thus, if dosage, format, *and* program type were put into the same analysis models, the question being answered would be whether there is some component of a program *not* represented in dosage or format that makes a difference in outcome. While an interesting question, it does not answer the more realistic question of which programs *as packages* of program components, length, and format show greater effects on school dropout. Therefore, the third model was designed to give some indication of the influence of dosage and format irrespective of program type. Below, we will examine the influence of the different program types, when treated as a package of formatting, dosage, and program components.

In general, programs delivered in classroom settings and those delivered in mixed or multiple settings (these typically had both school and community components) were significantly more effective than the programs in the reference group (i.e., those delivered strictly in community settings). The school-based programs not delivered during regular class time were neither more nor less effective than the other formats. Longer programs were not associated with better or worse outcomes, though programs with more frequent contacts tended to produce smaller effect sizes than those with less frequent contact (b = -.01; p < .10). Note, however, that the correlation matrix in Table 6.2 (Appendix I) indicates significant relationships between the program format, dosage, and program types. Some program types tended to be longer or shorter than others, and some were more typically delivered in certain types of settings. For example, school restructuring programs were nearly always delivered in classroom settings and tended to have more frequent contact than some of the other program types. Therefore, it is unknown whether it is the frequency of service, or the program types with which frequency is associated, that is the active ingredient.

Finally, we examine the influence of implementation quality, which was significantly associated with better treatment outcomes in Model III (b = .30; p < .05). Programs which had difficulties with implementation fidelity tended to show smaller effects on dropout than programs which indicated that no implementation problems were present, or did not explicitly identify any problems. While it is possible to imagine that research teams who found their programs to be less successful than expected

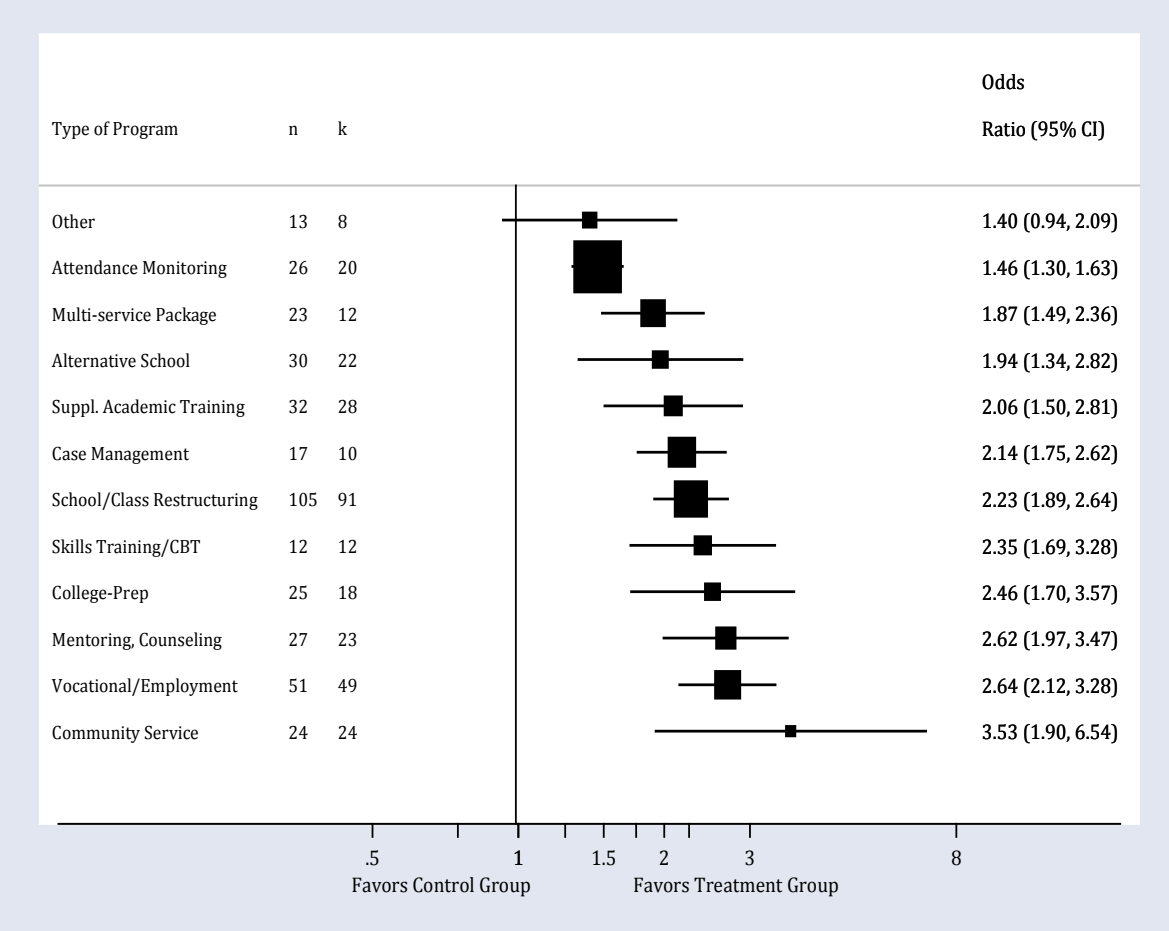
might be more likely to report complications with implementation of their programs, or that research teams who found significant treatment effects might spend less time in research reports on fidelity. However, we carefully review the supplementary research reports of eligible studies for any information related to implementation quality and believe that the finding here that implementation quality is important is not solely the result of research teams using implementation quality to explain null results.

3.3.6.1 *Effects of Different Program Types on School Dropout*

We now turn to an examination of the different program types represented in the review. To isolate the effects of program types, we wanted to remove the influence of study method, participant characteristics, and implementation quality, in order to see the effects of the different programs net of the influence of these other study characteristics. We did not, however, wish to remove the influence of the dosage and format variables, since those variables shared variance with the types of programs. Controlling for those variables would make it difficult to identify the effects of the programs as the packages of dosage and format that are typically implemented. For example, it would not make sense to remove the influence of classroom formatting in order to study the effects of school restructuring programs, which are nearly always implemented in classrooms. Therefore, we used Model II above, plus the implementation quality variable, to adjust the odds ratios for variation associated with those variables. Specifically, we produced the residuals from the regression model and added those to the predicted value associated with the mean values of all the variables in the model to obtain adjusted odds ratio estimates. This procedure artificially makes every study equal on all the variables in the model, thus allowing us to better see any remaining differences that might be associated with the different types of programs.

The adjusted mean odds ratios and confidence intervals for each of the program types are shown in Figure 3.1. Several findings are notable from the figure. First, the mean odds ratios for all of the programs except those in the “other” category are statistically significant, indicating that, on average, students in those programs have significantly lower dropout rates than students in the comparison conditions. The overlapping confidence intervals for most of the programs further indicate that no single program category stands out as significantly more or less effective than any other. However, while still producing significantly positive results, the attendance monitoring programs were significantly less successful than case management, school restructuring, skills training, college-oriented programs, mentoring/counseling, vocational programs, and community service. Dropout programs of most types shown in the figure tend to be effective at reducing school dropout, and, with the exception of the attendance monitoring programs, most produce about equal effects.

Figure 3.1: Mean Adjusted Odds Ratios for the General Dropout Programs, By Program Type



To better interpret the adjusted odds ratios shown in Figure 3.1, we have converted each of the mean odds ratios into percentage terms. These results are shown in Table 3.5. As above, we used the average dropout rate for control group participants of 21.1% as the baseline rate. For community service programs, the adjusted odds ratio of 3.53 translates into a dropout rate of about 7% for the intervention group, a rather large difference. The adjusted dropout rates for vocational programs, mentoring/counseling, and college-oriented programming all translate into less than 10% dropout after receiving the programs. Even the programs with the smallest odds ratios, such as attendance monitoring, translate into non-trivial changes in dropout rates, from 21.1% to about 15.5%.

Table 3.5: Mean Adjusted Odds Ratios and success rates By Program Type

	Mean OR	95% CI	n	k	Control group dropout rate	Program group dropout rate
Community service	3.53	1.90-6.54	24	24	21.1%	7.0%
Vocational training	2.64	2.12-3.28	51	49	21.1%	9.2%
Mentoring, counseling	2.62	1.97-3.47	27	23	21.1%	9.3%
College-oriented programming	2.46	1.70-3.57	25	18	21.1%	9.8%
Skills training, CBT	2.35	1.69-3.28	12	12	21.1%	10.2%
School, class restructuring	2.23	1.89-2.64	105	91	21.1%	10.7%
Case management	2.14	1.75-2.62	17	10	21.1%	11.1%
Supplemental academic services	2.06	1.50-2.81	32	28	21.1%	11.5%
Alternative school	1.94	1.34-2.82	30	22	21.1%	12.1%
Multiservice package	1.87	1.49-2.36	23	12	21.1%	12.5%
Attendance monitoring	1.46	1.30-1.63	26	20	21.1%	15.5%
Other	1.40	.94-2.09	13	8	21.1%	16.0%

3.4 DROPOUT PROGRAMS FOR TEEN PARENTS

3.4.1 Types of Dropout Programs for Teen Parents

We now turn to the results for the teen parent programs. There were 51 independent samples represented in the set of teen parent programs, all of them were 100% female. These samples generated a total of 119 odds ratios indexing differences between intervention and comparison groups after treatment. There were fewer types of programs for the teen parents, and they generally fell into two categories, as described in Table 3.6. The most common type of program for teen parents was a multi-service package. These programs generally included a wide range of services for the young mothers, including remedial education or GED preparation, vocational or other employment-related training, case management, and health and day care services. The other category of programs common for teen parents were the attendance monitoring and contingency programs – many of these programs provided financial incentives for teen mothers to return to school, or tied the receipt of welfare payments to school attendance contingencies (e.g., the LEAP study). Like the multi-service packages, the attendance monitoring with contingencies programs also tended to be quite comprehensive, and frequently included case management, transportation assistance, and child care services.

Table 3.6: Intervention Types–Teen Parent Programs

Intervention Type	Description	Number of samples (*k*)	Number of effect sizes (*n*)
Multi-service package	Large, comprehensive programs; often included academic, vocational, & case management, plus child care services, e.g., Project Direction, Project New Chance	34	54
Attendance monitoring & contingencies	Monitoring and services to increase attendance; most offer financial incentives and/or contingencies tied to welfare, e.g., the LEAP and Cal-Learn Programs	17	65

3.4.2 Methodological Characteristics

Table 3.7 presents the methodological, participant, and intervention characteristics for the dropout programs designed for teen parents. As with the general programs, a large proportion of the literature (about 80%) was unpublished, mainly in technical reports. All of the studies were conducted in the United States. Regarding methodological characteristics, 71% used random assignment, attrition averaged 14%, and 75% of the odds ratios were calculated using data that had been adjusted for covariates by the primary researchers. As with the general programs, most evaluators (75%) were independent of the programs they were evaluating.

3.4.3 Program Dosage, Format, and Location

Very few of the teen parent programs were delivered exclusively in school settings. Most of the programs were conducted in the community, or in multiple or mixed settings, as would be expected given the wide range of services generally provided by the programs. Treatment dosage was intensive, though program duration and frequency were shorter than for the general programs. Duration averaged about 53 weeks (compared to the general programs which averaged about 90 weeks). Frequency of program services varied across programs, with a group of 17 programs offering once a week services and another group of 17 programs offering daily services. Implementation quality was generally problematic. Only 7 studies had no apparent problems with implementation. The remaining 44 programs either evidenced clear problems or mentioned implementation issues that might have influenced study outcomes.

Table 3.7: Characteristics of Included Studies – Teen Parent Programs

Study Characteristics					
General Characteristics	*mean*	*sd*	Program Characteristics		
Publication year	1992	6	*Role of Evaluator*	*k*	*%*
			Delivered program	3	6
			Involved in planning, not		
Form of Publication	*k*	*%*	delivery	8	16
Journal article	10	20	Influential, but no direct role	2	4
Technical report	39	77	Independent evaluator	38	75
Dissertation	1	2			
Other	1	2	*Program Site*	*k*	*%*
			School classroom	3	6
Country of Study	*k*	*%*	School, but not classroom	2	4
United States	51	100	After school	0	0
			Community site	22	43
Method Characteristics			Mixed or multiple sites	24	47
Assignment Method	*k*	*%*			
Random (individual)	36	71	*Implementation Quality*	*k*	*%*
Random (group)	0	0	Clear problems	23	45
Matched	9	18	Possible problems	21	41
Nonrandom, nonmatched	6	12	No apparent problems	7	14
	mean	*sd*	*Program Dosage*	*mean*	*sd*
Attrition	14%	11%	Duration (weeks)	53	41
not reported in 1 study			Hours per week	13	10
					118
			Total contact hours	673	1
*Adjusted Data Used in OR**	*n*	*%*	*not reported in 16 studies*		
Yes	89	75			
No	30	25	*Frequency of Program*	*k*	*%*
			Less than weekly	6	13
Participant Characteristics	*mean*	*sd*	Once a week	17	37
Percent female	100	--	Twice a week	1	2
			3-4 times a week	5	11
	mean	*sd*	Daily (5 days/week)	17	37
Percent White	24	23	*not reported in 5 studies*		
Percent Black	56	31			
Percent Hispanic	19	25	Dependent Variables*	*n*	*%*
Percent other minority	3	4	Dropout	6	5
			Graduation	15	13
	mean	*sd*	Graduation or GED	55	46
			Enrolled (not dropped,		
Average age	18	1	graduated)	43	36
54% not in school					

Notes: Number of effect sizes (n = 119); Number of independent samples (k = 51)
**Reported at effect size level (n = 119)*

3.4.4 Participant Characteristics

The participants in the teen parent programs were somewhat older than the students in the general programs, with an average age of about 18. At the beginning of the studies, about 54% were not attending typical schools. All participants were female and ethnicity was diverse, with African American young women representing the largest ethnic group.

3.4.5 Synthesis of Results for Teen Parent Programs

Recall that the research studies reviewed here used multiple operationalizations of our outcome of interest: school dropout. In the analyses we conducted for the general programs, dummy variables were included to test and control for any differences associated with the different variations of outcome variables employed in the primary studies. These had little influence on the outcomes. However, for the teen parent programs, the different operationalizations of the outcome variables, especially those relating to school enrollment, were both associated with the size of the odds ratios and almost completely confounded with other important study characteristics. This circumstance made it difficult to isolate the influence of those study characteristics on the odds ratios when all of the odds ratios were analyzed together. We therefore split the odds ratios for the teen parent programs into two groups: one for the enrollment outcomes and a second for the outcomes representing dropout, graduation, and GED attainment. The random effects weighted mean odds ratios, confidence intervals, and heterogeneity statistics for the teen parent programs are shown in Table 3.8, separated by the two categories of outcomes—graduation/dropout and school enrollment. The mean odds ratio for the graduation outcomes was 1.83; it was statistically significant and indicated that program group participants had significantly higher graduation rates than those in control conditions. The mean odds ratio for the enrollment outcomes was smaller at 1.55, but also statistically significant.

Table 3.8: Mean Odds Ratio Effect Sizes and Heterogeneity Statistics, by Outcome Type

	Mean OR	*95% CI*	*n*	*k*	Q_E	τ^2
Teen Parent Programs						
Graduation/Dropout Outcomes	1.83	1.54, 2.18	86	48	211	.35
Enrollment Outcomes	1.55	1.15, 2.09	33	15	35	.16

Notes: Mean effect sizes and confidence intervals estimated with robust standard errors to account for dependent effect sizes estimates. N– number of effect sizes; K – number of unique samples; Q_E – weighted sum of squares heterogeneity statistic that accounts for dependent effect size estimates, τ^2 – method of moments between studies variance component that accounts for dependent effect size estimates.

For interpretation, we can convert these odds ratios back into proportions. The average graduation rate for the young women in comparison groups was 26%.[3] This is extremely low and illustrates the high risk for dropout of the sample of young mothers in the studies we reviewed. The corresponding graduation rate for young mothers in the intervention programs was 39%, a clear improvement over the 26% graduation rate for control group participants. For the enrollment outcomes, the average enrollment rate for the comparison mothers was 33%. The odds ratio of 1.55 for these studies translates into an enrollment rate of about 43%, a ten percentage point improvement for the intervention groups.

The heterogeneity statistics for the graduation and dropout outcomes for the teen parent programs suggest that there is more variability in the distribution of odds ratios than would be expected from sampling error, warranting further exploration of the relationship of the various study characteristics to the odds ratios is warranted. The heterogeneity statistics for the enrollment outcomes are also statistically significant. We now turn to a presentation of the moderator analyses we performed for the teen parent programs, focusing on the set of 86 odds ratios indexing graduation or dropout. Moderator analyses were not possible with the enrollment outcomes because the key moderator variables did not vary within the set of programs providing enrollment outcomes. That is, across the programs providing enrollment outcomes, all or nearly all used the same method of assignment (random), were about the same duration and frequency, had implementation problems, etc. Without variability in the moderator variables, moderator analyses were not possible. A forest plot for the individual teen parent programs providing enrollment outcomes is presented in Figure 7.2 in Appendix II.

3.4.6 Moderator Analysis for the Teen Parent Programs

Moderator analyses were performed for the 86 odds ratios (48 studies) of teen parent programs with dropout or graduation outcomes. As with the general programs, we present three models. Model I included the method characteristics, Model II added the participant characteristics, and Model III included dosage and implementation quality. The intent behind the three models was to provide some assessment of the separate influence of method, participant, and dosage characteristics on the outcomes. Because most programs are packages of program elements and dosage, we elected to examine the dosage variables on their own first. Below, we will examine the results for the two different program types: attendance monitoring and contingencies vs. the multi-service packages.

[3] Note that we used the average graduation rate here instead of the average dropout rate we used for the conversions for the general programs because graduate was the most common operationalization of the outcome for these programs.

Table 3.9: Meta-Regression Results Predicting Logged Odds Ratio Effect Sizes–Teen Parent Programs

	Model I			Model II			Model III		
	B	se	95% CI	b	se	95% CI	b	se	95% CI
Methodological Characteristics									
Random assignment (1=yes)	-.89*	.20	-1.30, -.49	-1.28*	.23	-1.75, -.82	-1.12*	.42	-1.97, -.26
Matched groups design (1=yes)	-1.26*	.54	-2.35, -.17	-1.30*	.48	-2.28, -.32	-1.57*	.47	-2.53, -.62
Attrition (%)	-.34	.77	-1.91, 1.22	-1.07	.84	-2.76, .62	-.37	.85	-2.09, 1.35
Mean group equivalence OR	.23	.29	-.37, .82	.25	.43	-.61, 1.11	-.11	.33	-.78, .56
Adjusted data for OR calc (1=yes)	.03	.30	-.57, .64	.03	.30	-.57, .63	.39	.34	-.31, 1.08
Independent evaluator (1-4)	-.34*	.15	-.64, -.05	-.50*	.16	-.82, .19	-.34†	.18	-.71, .03
Participant Characteristics									
Percent White				-.45	.59	-1.65, .75	-.03	.52	-1.09, 1.03
Average age				.31*	.15	.01, .61	-.08	.21	-.50, .33
Program Characteristics									
Program duration (in weeks)							-.00	.00	-.01, .00
Hours of service per week							.01	.01	-.01, .03
Implementation quality (1-3)							.53*	.15	.24, .83

Notes: Mixed effects meta-regression models estimated with robust standard errors to account for dependent effect sizes estimates. b – unstandardized regression coefficient.

† $p < .10$. * $p < .05$.

Methodological variables were influential in all three meta-regression models of the teen parent programs, with assignment method and the role of the evaluator significantly associated with the odds ratios. Both random assignment and matched designs produced significantly smaller odds ratios than the non-random, non-matched research designs included in the reference group. Though not consistently significant in the meta-regression models for the general dropout programs, the direction of the relationship was the same for general programs and teen parent programs. This is likely a reflection of the fact that the non-random, non-matched designs may have had less equivalent groups at the baseline than the random assignment and matched designs. If those differences favored the control groups (i.e., the intervention groups were at higher risk for dropping out), we would expect the post-treatment effect sizes to further accentuate any initial group differences. Independent evaluators were also significantly associated with smaller odds ratios for the teen parent programs; research conducted by independent evaluators found dropout programs for teen parents to be less effective than research conducted by evaluators more closely involved in the intervention. The influence of independent evaluators on the effects of general dropout programs was not significant, though independent evaluators were similarly associated with smaller treatment effects (i.e.,

the direction of the relationship was the same for teen parent and general programs).

With regard to the participant characteristics, ethnicity was not reliably associated with the odds ratios for teen parent programs. In general, the dropout programs for young mothers appeared equally effective across majority and minority women. When we look at Model II, the older students tended to exhibit better outcomes than younger students. This relationship is no longer significant when we examine Model III, even though age is not significantly correlated with the dosage and implementation variables (see Table 6.3 in Appendix I).

Treatment duration and frequency of service were not significantly associated with larger or smaller odds ratios; longer or more frequent programs did not appear to produce better results than shorter or less frequent programs. Implementation quality, however, was significantly associated with outcomes. Programs that experienced implementation difficulties tended to have smaller effects on school completion than programs that did not appear to have implementation problems.

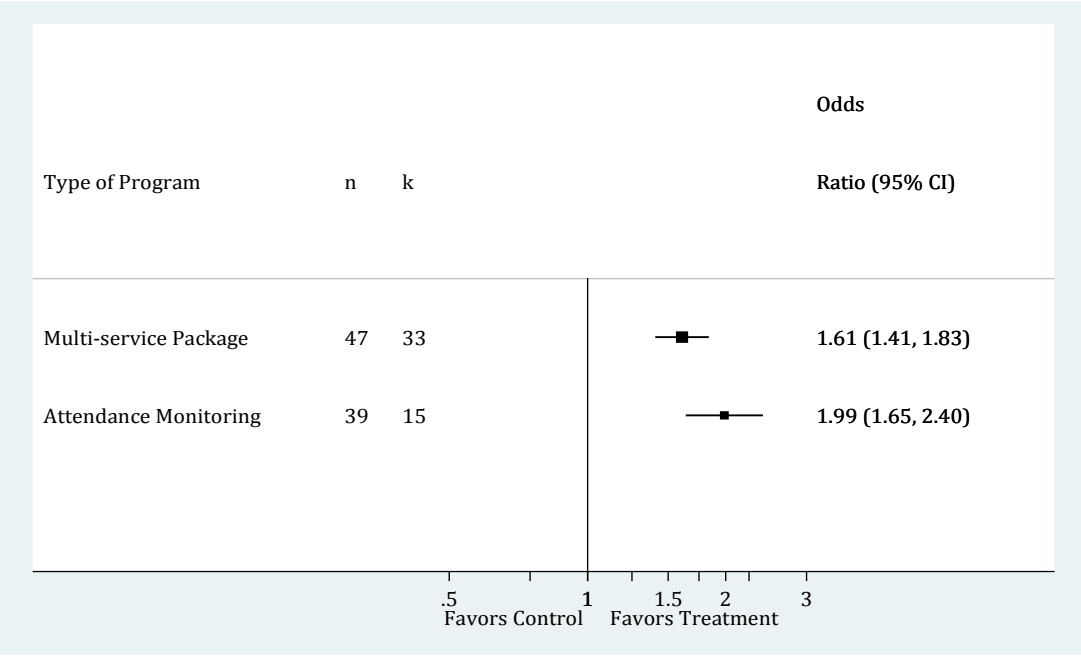
3.4.6.1 *Effects of Different Teen Parent Programs on School Dropout*

The next step in our analysis of the dropout programs for teen parents was to produce odds ratios adjusted for some of the moderators we identified in the meta-regression models. The purpose of these adjustments was to control for the method characteristics, subject characteristics, and implementation quality in order to better isolate any differences between the two program types: attendance/monitoring plus contingencies and multi-service packages. Because treatment dosage is confounded with program type, we did not control for duration or service frequency when performing the adjustments. As above, we produced the residuals from the regression model and added those to the predicted value associated with the mean values of all the variables in the model to obtain adjusted odds ratio estimates. The mean adjusted odds ratios and confidence intervals for the two program types are shown in Figure 3.2 and the conversions to percentages are shown in Table 3.10.

Table 3.10: Mean Adjusted Odds Ratios And Success Rates by Teen Parent Program Type

	Mean OR	95% CI	N	K	CT	TX
Multi-service Package	1.61	1.41, 1.83	47	33	26%	36%
Attendance monitoring w/ contingencies	1.99	1.65, 2.40	39	15	26%	41%

Figure 3.2: Mean Adjusted Odds Ratios for the Teen Parent Dropout Programs



The young mothers in both types of programs exhibited significantly higher graduation rates than the young mothers in the comparison conditions. The overlapping confidence intervals for the two program types further indicate that there were no significant differences in effectiveness between the two program types. While the adjusted mean odds ratio for the attendance monitoring programs is higher than that for the multi-service packages, the difference is not statistically significant. Furthermore, using an average graduation rate of 26% for the comparison mothers, the adjusted odds ratios in Table 3.10 translate into large improvements in graduation rates for both types of teen parent programs (36% for the multi-service packages and 41% for attendance monitoring and contingencies).

3.5 PUBLICATION BIAS ANALYSIS

Given the large number of unpublished studies of dropout interventions (indeed, the majority of studies included in this review were unpublished), we did not anticipate that there would be any substantial publication bias manifested in this review. However, many of the unpublished technical reports included in the review involved large, multi-site studies. Thus, while we may have obtained a large number of unpublished studies with null or small effects, it is possible that the studies in the review do not include smaller studies with null effects that were never submitted for publication. There are a number of methods for assessing bias associated with missing small studies (see, for example, Rothstein, Sutton, & Borenstein, 2005), including producing funnel plots and conducting Egger's regression-based test. The Egger's test, however, has inflated Type I error rates with odds ratio effect sizes, particularly when mean effect sizes are large, like those reported here (Peters, Sutton, Jones, Abrams, & Rushton, 2006). We therefore elected to assess publication bias with a visual analysis of funnel plots, and used the regression based

test for small study effects recommended by Peters and colleagues (2006) for odds ratio effect sizes.

Figures 3.3 and 3.4 show the funnel plots for the general and teen parent dropout programs, respectively. As shown in Figure 3.3, the funnel plot for general programs is reasonably symmetric around the fixed effect mean logged odds ratios. Most notable, however, is that most of the general program studies had large sample sizes (i.e., small standard errors), and thus there were very few small sample studies in general, regardless of effect size strength/direction.

Figure 3.3: Funnel Plot for the General Dropout Programs

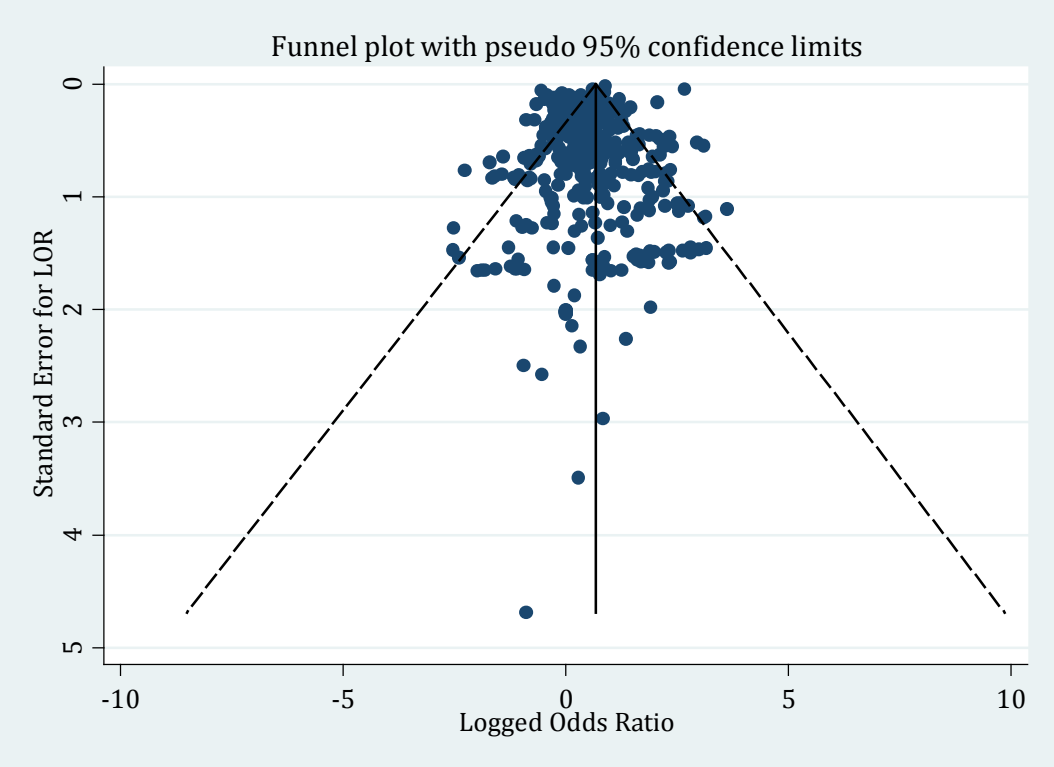
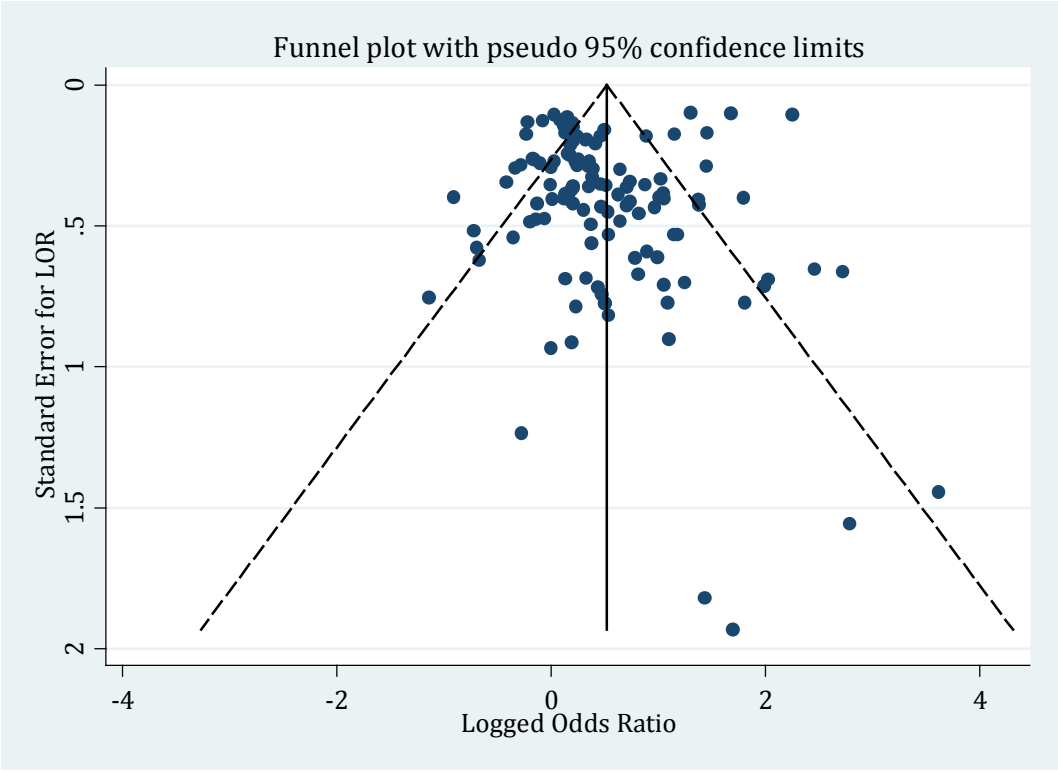


Figure 3.4: Funnel Plot for the Teen Parent Dropout Programs



The funnel plot in Figure 3.4 for teen parent programs appears somewhat less symmetric, however, indicating a possible absence of small sample studies with null or negative findings (i.e., odds ratios close to or less than zero). Again, however, there were relatively few small sample studies with positive or negative results.

Because interpretation of funnel plot asymmetry is inherently subjective, we also conducted a regression-based test recommended by Peters and colleagues (2006) for use with logged odds ratios. This test is a simple weighted linear regression which models the logged odds ratio against the inverse of the study sample size (in contrast to the standard error used in the Egger's test). Coefficients from the regression test are shown in Table 3.11 for both the general programs and the teen parent programs. The Peters' test was significant for the general dropout programs (b = 16.80; p = .001), but not for the teen parent programs (b = 6.13; p = .771). This indicates the possibility of small study effects in the general dropout prevention programs. Although we attempted to minimize the possibility of publication bias by conducting an extensive grey literature search, these results indicate the possibility of small study effects in the meta-analysis for general dropout programs. These "small study effects" could be due to reporting biases such as selective outcome reporting, spurious inflation of effects in smaller studies due to lower methodological quality, true heterogeneity in effect sizes for studies with different sample sizes, artifacts due to sampling variation, or purely chance (Harbord, Harris, & Sterne, 2009), and unfortunately it is impossible to determine the reason for the observed small study effects.

Table 3.11: Results of Peters' regression Test for Small Study Effects, By Type of Program

	b	se	t	p	n
General Programs	16.80	18.62	3.32	.001	385
Teen Parent Programs	6.13	20.98	.21	.771	119

Although we cannot test for the actual presence or absence of publication bias, results from Peters' test do provide some indication that publication bias may be a concern in the present literature of general dropout programs. As such, it is plausible that the mean effect size estimates for the general programs may be over-estimated if there are small sample studies with null or negative findings that may be omitted from the current analytic sample. It should be noted, however, that additional exploratory trim and fill analyses (Duval & Tweedie, 2000) yielded no trimmed/filled effect sizes, thus suggesting that even if small study effects may be an issue for the general dropout programs, they are not likely to have much substantive impact on the estimated mean effect sizes reported in this review.

3.6 SENSITIVITY ANALYSIS

Because missing data for the moderator variables were imputed for some studies, we conducted a series of sensitivity analyses to explore whether the imputed variables may have influenced the results. For each of the variables that contained imputed data, we re-ran Model III twice, first using the reduced dataset that only included the data that were coded from the primary studies (i.e., we dropped the cases with missing values on that variable); the second model dropped the variable entirely; thus, it used complete cases but removed the influence of that particular imputed variable. We performed these analyses for both the general and the teen parent programs separately. For both sets of programs, the imputed variables did not appreciably change the results; that is, no relationship between any moderator variable and the odds ratios was altered either in direction or in statistical significance.

4 Discussion

4.1 SUMMARY OF EVIDENCE

One hundred fifty-two (152) general dropout programs and 15 dropout programs for teen parents were reviewed. The random effects mean odds ratio for the general programs was 1.72. Using an average control group dropout rate of 21.1%, the mean odds ratio for the general programs translates into a dropout rate of 13% for the program groups. The mean odds ratios for the teen parent programs were 1.83 for graduation and dropout outcomes and 1.55 for school enrollment outcomes. For the teen parent programs, we used the average comparison group graduation rate of 26% to translate the odds ratio of 1.83 into a graduation rate of 39% for the program group. For school enrollment outcomes for the teen parent programs, the average enrollment rate of 33% for comparison groups translates to an enrollment rate of 43% for program participants. These results clearly indicate substantial improvements in school completion and dropout rates for primary and secondary students participating in prevention and intervention programs and for teen parents participating in programs designed to increase school completion.

4.1.1 General Dropout Programs

Although most of the general programs produced positive effects on dropout (75% of the effect sizes were positive), there was considerable variability across the programs in overall effectiveness. Moderator analysis of the general programs revealed that while study design per se was not associated with larger or smaller effects, variables indexing initial equivalence between intervention and comparison groups and whether adjusted data were used in the odds ratio calculations were both significantly associated with the odds ratios. Studies with larger pre-treatment differences between groups tended to exhibit larger post-treatment differences between the groups. Furthermore, post-treatment odds ratios that were not adjusted for covariates were also associated with larger differences between the groups after treatment. While the impact of evaluation methods on study results may not be an issue most practitioners or policymakers consider in making programming decisions, the methods by which programs are evaluated can exert a substantial influence on how effective those programs appear in the final results. How the methods might influence study results may not always be clear in individual evaluation reports. By compiling large numbers of studies of relevant programs and

controlling for the influence of study method, as we have done in this systematic review, a more accurate picture of results expected from particular program strategies can emerge. Both the meta-regression models and the adjusted mean odds ratios shown in Figure 3.1 and Table 3.5 provide estimates of the influence of the program characteristics (e.g., implementation quality) and program type *net* of the methodological characteristics. Thus, while the study designs with more equivalent groups at baseline tended to produce smaller post-treatment effects, our analyses allowed us to estimate treatment effects that adjust for the influence of these methodological characteristics.

Overall, participant demographics had minimal influence on dropout program effectiveness. The gender and racial composition of samples had no significant influence on general programs' effectiveness, although effects were somewhat larger for older students ($b = .04$; $p < .10$). For instance, the predicted mean odds ratio for 15 year old students was 1.75 versus 1.89 for 17 year old students participating in general programs. In terms of percent dropout, these differences translate into about a 1% difference in dropout rates for 15 year olds vs. 17 year olds. While practitioners may still wish to tailor programs to particular racial or ethnic groups to encourage student engagement, or try different strategies with different age groups who may have different academic needs, the findings here suggest that such tailoring isn't necessary for programs to be effective at reducing dropout.

Consistent with expectations from the prior literature, program setting was also associated with dropout outcomes. Classroom-based programs and the mixed-setting programs produced significantly larger reductions in dropout than the community-based programs. Keep in mind, however, that even the community-based programs produced positive outcomes; they were just smaller than those for the school-based and mixed-setting programs. Furthermore, setting overlaps considerably with program type, so it is unclear whether it is the setting or the actual program types (in their settings) that are driving this finding. Less frequent programs were also more effective ($p < .10$). The programs with more hours per week, though still producing positive effects, tended to show smaller reductions in dropout. Although this finding may initially appear counterintuitive, the more frequent programs may have experienced more variability in program delivery or receipt of services, resulting in the smaller, but still positive, outcomes. Finally, implementation quality was highly significant in the meta-regression models. Programs experiencing implementation problems tended to exhibit significantly smaller reductions in dropout.

To examine the comparative effectiveness across the different types of programs, we produced adjusted odds ratios that controlled for key moderator variables. The results of that analysis showed that all types of dropout programs (except the few programs in the "other" category) were effective and that most dropout programs were about equally effective.

4.1.2 Dropout Programs for Teen Parents

The results were similar for the programs designed for teen parents. Overall, most studies had positive outcomes, though moderator analyses identified influential study characteristics. Study design was associated with the odds ratio effect sizes, with the random assignment and matched designs generally producing smaller odds ratios than the non-random, non-matched designs. Furthermore, independent evaluators were associated with smaller odds ratios. Among the program characteristics, implementation quality was very important, net of the methodological moderators. Studies experiencing implementation difficulties produced smaller odds ratios.

There were only two types of programs for teen parents, and both generally offered a range of comprehensive services. The multi-service packages frequently included academic and vocational components, plus case management and child care services. The attendance monitoring programs focused on providing financial contingencies for continued school enrollment and attendance, but also offered a range of services to the young mothers. When the odds ratios were adjusted to equalize the influence of the methodological and participant variables, both types of programs were effective at increasing graduation rates, and neither strategy was significantly more effective than the other.

4.2 LIMITATIONS

There are several limitations to the current review that should be acknowledged. First, results from publication bias analysis indicated the possibility of small study bias in results for the general program results. There is a small chance that the results from the meta-analysis may be over-estimating the effects of general programs on dropout outcomes if the sample is missing a theoretical subset of studies with small sample sizes and null or negative findings. We attempted to minimize the possibility of publication bias by conducting an extensive literature search, with particular emphasis on locating grey literature – and indeed, the large number of unpublished technical reports and dissertations in the sample attest to this search effort. Indeed, this review includes many more studies than any previous review on the topic. As previously noted, however, many of the unpublished technical reports were actually large multi-site studies, and very few small sample size studies were included in the meta-analysis. Although exploratory trim and fill analysis provided some assurance that any small study bias is likely to have minimal influence on the substantive interpretation of the meta-analysis results, all conclusions from the current review should be sensitive to the possibility of small study bias.

It should also be acknowledged that the meta-analyses in this review synthesized effect sizes from studies using both experimental and quasi-experimental research designs, thus introducing a risk of bias associated with any lower quality quasi-

experimental research studies. We attempted to minimize any such bias by requiring any eligible non-randomized studies to include baseline pre-test or group equivalence information that could then be statistically adjusted for in the final meta-regression models used to estimate the comparative effectiveness of different program types. Interestingly, the random assignment studies were more likely to provide adjusted data, while the different research designs did not appear to have larger or smaller group equivalence effect sizes (something we might expect if random assignment designs tended to produce more equivalent groups). While the random assignment designs would technically equate the treatment and comparison groups on any variable (whether measured or not), the small relationship of research design to group equivalence suggests that any of the research designs could have experienced non-equivalence at baseline (and indeed they did).

The geographic setting of the programs we reviewed was not part of our coding scheme, though there may be differences in the types of program strategies employed in rural and urban settings, and in the effectiveness of those strategies across different settings. However, our finding that there were no differences in program effectiveness associated with the racial/ethnic mix of the samples has bearing here. Studies with high minority proportions (which also tended to be more urban) were not any more or less effective than studies with lower minority proportions. However, the issue of whether different programs are more effective in some geographic settings than others warrants further examination.

One aspect of program implementation was rarely (if ever) mentioned in the studies we reviewed – leadership. Whether programs with strong champions are better implemented, is an important question for additional research. It is probable that a good leader may be a critical component for program effectiveness, especially through encouraging better implementation of the program components, but we were unable to explore this question in the current review.

Another limitation of this review was the inability to synthesize cost benefit information for different prevention and intervention programs. Given the findings that no program type consistently outperformed others, along with the importance of implementation quality, it would have been desirable to summarize the different costs of programs, and therefore identify what programs may be most cost effective in which environments. Unfortunately, the summarized literature did not consistently report cost information and thus it was not possible to synthesize this information.

4.3 CONCLUSIONS

The most notable finding from our comprehensive systematic review of dropout programs was that no single prevention or intervention strategy stood out as better than any other. We believe this finding has particular practical significance, especially when taken together with the finding on the importance of

implementation quality for both general and teen parent programs. School administrators and policymakers have a great variety of choices available to them for dropout prevention and intervention programs. The results from this systematic review suggest that the particular program strategy chosen makes less of a difference in eventual outcome than selecting a strategy that can be implemented successfully by the school or agency. Unfortunately, cost information was rarely provided in the studies included in this review, but our findings suggest that decision makers may be better off considering the fit of a program with their setting and staff, and the cost of a program, than in selecting a particular or popular strategy. Once programs are in place, focusing on implementation quality is critical. Keep in mind, however, that most of the programs reviewed here were relatively intensive (occurring over significant time and involving considerable changes in the educational settings in which they were implemented). It would be unwise to conclude that less intensive, much less frequent or much shorter programs (which did not turn up often in the literature and were therefore not part of this review) would be as effective as those reviewed here.

Looking ahead, delving deeper into some of the other study outcomes that were coded for the dropout programs we reviewed, including attendance, engagement, school performance, and school behavior, may provide more clues to the ways in which dropout interventions might be effective. Furthermore, the large body of research we discovered that focuses mainly on attendance may also provide more understanding of the relationship between attendance and dropout.

5 References

Altonji, J. G., & Mansfield, R. (2010). *The contribution of family, school and community characteristics to inequality in education and labor market outcomes.* Unpublished manuscript, Yale University. Retrieved April 25, 2011 from http://www.econ.yale.edu/~rkm24/Altonji%20and%20Mansfield%20%282010%29.pdf

Arneberg, M. W. (2009). *Completion rates in upper secondary education.* Background Paper for the Informal Meeting of OECD Ministers of Education, Oslo 9th and 10th of June 2009. Obtained directly from the author.

Canadian Council on Learning. (2010). *2010 Composite learning index indicator fact sheets: High-school dropout rate.* Retrieved June 14, 2010 from http://www.ccl-cca.ca/pdfs/CLI/2010/2_HighSchoolDropoutrate.pdf

Cataldi, E. F., Laird, J., & KewalRamani, A. (2009). *High school dropout and completion rates in the United States: 2007* (NCES 2009-064). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved Jan 26, 2010 from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2009064.

Dalton, B., Glennie, E., & Ingels, S. J. (2009). *Late high school dropouts: Characteristics, experiences, and changes across cohorts* (NCES 2009-307). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Duval, S., & Tweedie, R. (2000). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98.

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology*, Vol. 2. (pp. 87-114). Hoboken, NJ: John Wiley & Sons, Inc.

Hammond, C., Linton, D., Smink, J., & Drew, S. (2007). *Dropout risk factors and exemplary programs: A technical report.* Clemson, SC: National Dropout Prevention Center/Network and Alexandria, VA: Communities In Schools, Inc.

Harbord, R. M., Harris, R. J., & Sterne, J. A. C. (2009). Updated tests for small-study effects in meta-analyses. *Stata Journal, 9(2)*, 197-210.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107-128.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics, 32,* 341-370.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39-65.

ICF International & National Dropout Prevention Center/Network. (2008). *Best practices in dropout prevention.* Fairfax, VA: ICF International.

Klima, T., Miller, M., & Nunlist, C. (2009). *What works? Targeted truancy and dropout programs in middle and high school.* Olympia: Washington State Institute for Public Policy, No. 09-06-2201. [http://www.wsipp.wa.gov/pub.asp?docid=09-06-2201]

Lehr, C. A., Hansen, A., Sinclair, M. F., & Christenson, S. L. (2003). Moving beyond dropout toward school completion: An integrative review of data-based interventions. *School Psychology Review*, 32, 342-364.

Long, D., Gueron, J. M., Wood, R. G., Fisher, R., & Fellerath, V. (1996). *LEAP: Three-year impacts of Ohio's welfare initiative to improve school attendance among teenage parents*. New York: Manpower Demonstration Research Corporation.

National Dropout Prevention Center/Network. (2009). *Economic impacts of dropouts*. Retrieved Jan 26, 2010 from http://www.dropoutprevention.org/ndpcdefault.htm

Organisation for Economic Co-operation and Development. (2009). *Education at a Glance 2009: OECD Indicators*. Paris, France: OECD Publishing. [Retrieved Jan 26, 2010 from http://www.oecd.org/edu/eag2009].

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *JAMA, 295,* 676-680.

Perper, K., Peterson, K., & Manlove, J. (2010). *Diploma attainment among teen mothers.* Child Trends, Fact Sheet Publication #2010-01: Washington, DC.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention, assessment and adjustments.* West Sussex, England: John Wiley & Sons, Ltd.

Sum, A., Khatiwada, I., & McLaughlin, J. (2009). The consequences of dropping out of high school. *Center for Labor Market Studies Publications.* Paper 23. http://hdl.handle.net/2047/d20000596

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

U.S. Department of Education What Works Clearinghouse. *Topic report on dropout prevention*. [http://ies.ed.gov/ncee/wwc/reports/dropout/topic/#top]

6 Appendix I: Additional Tables

6.1 NUMBER OF REPORTS EXCLUDED, BY SPECIFIC EXCLUSION CRITERIA

Inclusion Criteria	Number Excluded
Intervention A school-based or affiliated psychological, educational, or behavioral prevention/intervention program that involves actions performed with the expectation that they will have beneficial effects on student recipients. Community-based programs that are explicitly presented as dropout prevention/intervention programs.	317
Population Intervention directed toward school-aged youth (ages 4-18) or recent dropouts between the ages of 18-22 for programs explicitly oriented toward secondary school completion or the equivalent.	2
Population General population samples of school-age children. Samples consisting exclusively of specialized populations, such as students with mental disabilities or other special needs, are not eligible.	9
Research Design Experimental or quasi-experimental design; must involve comparison of treatment and control conditions.	579
Research Design Students are randomly assigned or non-randomly assigned with matching, statistical controls, or evidence of initial equivalence on key risk variables or student characteristics.	174
Outcomes Intervention effects are assessed on at least one outcome variable that represents school completion or dropout, or is a close proxy measure or recognized precursor for dropout.	283
Outcomes The report presents an evaluation of a program. Reports providing solely program descriptions are not eligible.	146

Inclusion Criteria	Number Excluded
Study Characteristic Date of publication or reporting of the study is no later than 1985. Research must be conducted in 1980 or later.	59
Study Characteristic The study must report sufficient quantitative data to compute an effect size on an eligible outcome and the variables involved in the effect size must have a known direction of scoring.	64

6.2 BIVARIATE CORRELATION MATRIX OF STUDY CHARACTERISTICS—GENERAL PROGRAMS

Study Characteristics	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	.26	.27	.28	.29
1. Random assignment design (1=yes)	-.60	.25	.06	.33	-.13	-.05	.21	-.08	.09	-.02	.03	-.14	-.01	.01	-.11	-.18	-.28	.04	-.24	.12	.10	.03	.08	.20	.01	-.20	-.07	.16
2. Matched groups design (1=yes)		-.07	-.10	-.29	.22	-.10	-.20	.16	-.13	.07	.02	.04	.08	-.03	.11	.12	.01	-.13	.26	-.14	-.01	-.06	-.19	-.10	-.07	.32	.16	-.11
3. Attrition (%)			.04	.15	-.25	.18	.10	-.01	.07	-.04	-.11	-.25	-.01	.15	.26	-.19	-.15	.01	-.13	-.08	.10	-.02	-.10	.14	.02	-.12	.17	.14
4. Mean pretest odds ratio from study				.03	-.10	.11	.05	-.11	-.07	.21	-.29	.01	.02	-.01	-.14	.05	-.06	-.01	.04	.07	-.12	-.04	.07	-.06	-.02	.08	-.04	-.05
5. Adjusted data used for OR calc (1=yes)					-.21	.16	.16	.13	.04	-.16	-.12	.02	-.02	.09	.13	.07	-.08	-.02	.09	.08	-.16	.17	-.05	.29	-.10	-.12	-.13	.08
6. Dropout outcome						-.67	-.40	.04	.07	.09	.13	.16	.17	-.30	-.29	.10	.06	.10	.16	.03	.02	.05	-.13	-.20	-.07	.20	-.08	-.19
7. Graduation outcome							-.16	-.05	-.05	-.07	-.25	.05	-.19	.18	.30	-.01	.00	-.09	.00	-.04	-.07	-.06	.07	.10	.12	-.14	.08	.03
8. Graduation/GED outcome								.04	-.04	.00	.14	-.29	-.02	.25	.09	-.15	-.08	.01	-.19	-.02	-.05	-.06	.23	.25	.03	-.08	-.09	.04
9. Role of evaluator									-.09	-.03	.23	-.05	-.07	.08	.17	.00	-.13	.09	.08	-.10	-.21	-.24	.17	.07	.03	.12	-.03	.10
10. Percent male										.03	-.04	.14	-.11	-.08	.05	.10	-.08	.03	.08	-.04	.10	.02	.10	.02	-.07	-.28	-.06	-.02
11. Percent White											.01	.05	.09	-.22	-.26	-.03	-.03	.06	-.01	.01	-.03	-.04	.00	-.08	.01	.14	-.05	-.09
12. Average age of sample												-.04	-.06	-.04	-.31	.09	-.04	.05	-.04	.17	-.03	-.02	.17	-.13	-.03	-.02	-.04	-.02
13. Delivered in class (1=yes)													-.53	-.42	-.06	.55	-.08	-.13	.50	.14	-.12	.17	-.01	-.23	-.17	-.10	-.21	-.22
14. Delivered at school, not class time (1=yes)														-.20	-.23	-.42	.21	.27	-.31	-.15	.12	-.09	-.06	-.01	.00	.30	.04	.15
15. Delivered in mixed settings (1=yes)															.49	-.19	-.14	-.12	-.16	.06	-.11	-.07	.02	.43	.02	-.10	.22	.03
16. Program duration (weeks)																-.01	-.15	-.18	.08	-.11	.00	-.18	-.04	.33	.09	-.19	.13	.01
17. Program hours per week																	-.16	-.06	.63	.30	-.26	-.14	-.16	-.22	-.13	-.26	-.24	-.06
18. Implementation quality (higher=better)																		.07	-.12	.01	.11	-.02	.09	-.26	-.01	.21	-.02	-.14

Study Characteristics	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	.26	.27	.28	.29
19. Supplemental academic services																			-.18	-.09	-.08	-.05	-.12	-.06	-.08	-.08	-.08	-.08
20. School, class restructuring																				-.18	-.16	-.11	-.24	-.13	-.16	-.16	-.17	-.15
21. Alternative school																					-.08	-.05	-.11	-.06	-.08	-.07	-.08	-.07
22. Attendance monitoring																						-.05	-.11	-.06	-.07	-.07	-.07	-.07
23. Skills training, CBT																							-.07	-.04	-.05	-.05	-.05	-.05
24. Vocational program																								-.08	-.10	-.10	-.11	-.10
25. Case management																									-.06	-.06	-.06	-.05
26. College-oriented program																										-.07	-.07	-.07
27. Community service																											-.07	-.07
28. Mentoring, counseling																												-.07
29. Multiservice package																												

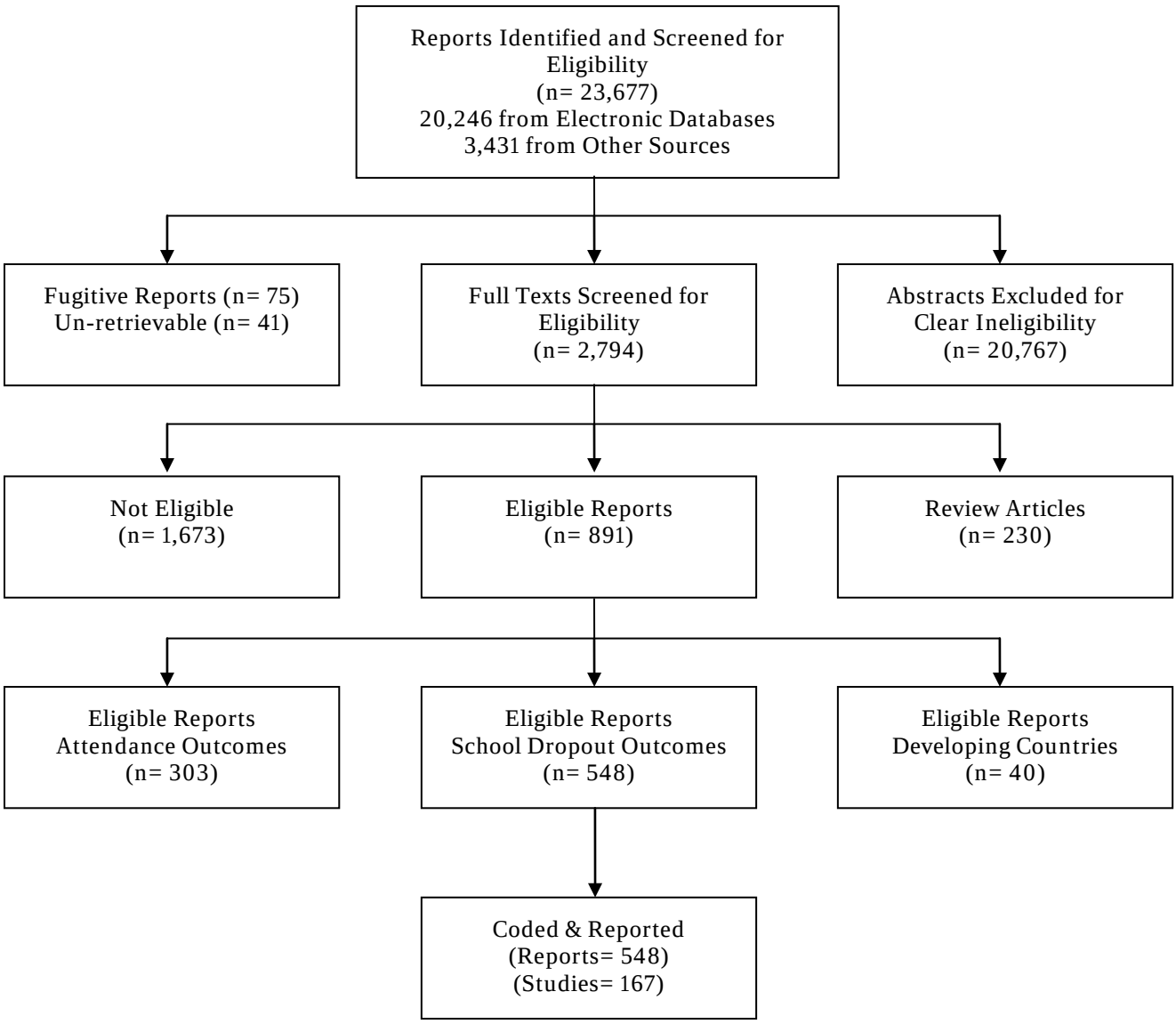
Notes: Correlations greater than .10 or less than -.10 are statistically significant ($p < .05$; $n = 379$).

6.3 BIVARIATE CORRELATION MATRIX OF STUDY CHARACTERISTICS—TEEN PARENT PROGRAMS WITH DROPOUT OR GRADUATION OUTCOMES

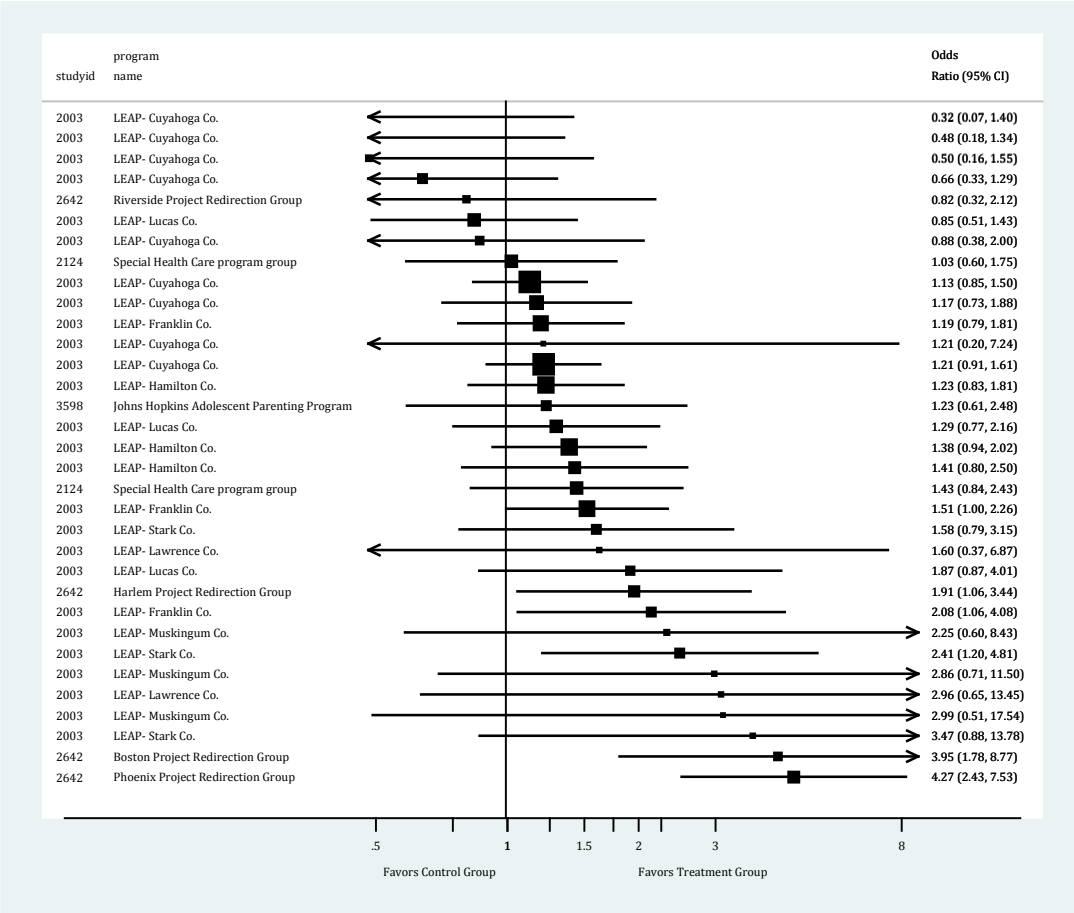
Study Characteristics	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Random assignment design (1=yes)	-.68.	-.28.	.12.	.56.	.61.	.29.	.70.	.20.	-.03.	-.34.	.31.
2. Matched groups design (1=yes)		.17	-.25	-.37	-.80	-.30	-.64	-.13	.14	.47	-.42
3. Attrition (%)			.04	-.51	-.24	-.46	.02	.31	.24	.33	-.54
4. Mean pretest odds ratio from study				.25	-.01	.29	.19	-.06	-.48	-.12	-.04
5. Adjusted data used for OR calc (1=yes)					.41	.51	.30	-.28	-.51	-.61	.47
6. Role of evaluator						.35	.68	-.02	.07	-.46	.43
7. Percent White							.14	-.35	-.29	-.45	.48
8. Average age of sample								-.07	.11	-.01	-.11
9. Program duration (weeks)									.03	.02	.01
10. Program hours per week										.50	-.35
11. Implementation quality (higher=better)											-.77
12. Attendance monitoring w/ contingencies prog. (1) vs. multi-service packages (0)											1.0

7 Appendix II: Additional Figures

7.1 STUDY IDENTIFICATION FLOW DIAGRAM



7.2 FOREST PLOT FOR TEEN PARENT PROGRAMS PROVIDING SCHOOL ENROLLMENT OUTCOMES



Notes: Studies were allowed to contribute more than one effect size estimate either by including different operationalizations of the outcome variables or by including multiple non-overlapping subgroups of participants (e.g., LEAP Counties). For instance, effect size estimates could be calculated separately for White girls, Black girls, and Hispanic girls.